

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:25:43 ; Search time 1.09687 Seconds
(without alignments)
3164.197 Million cell updates/sec

Title: US-09-743-684A-19
Perfect score: 63
Sequence: 1 GDVCQDCIQMW 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	50	79.4	520	13 Q8UVZ4	Q8uvz4 brachydanio
2	50	79.4	522	13 Q9DG82	Q9dg82 brachydanio
3	47	74.6	554	11 Q8BFO1	Q8bfq1 mus musculus
4	45	71.4	402	11 Q8CLN0	Q8cln0 mus musculus
5	45	71.4	449	11 Q8BJV5	Q8bjv5 mus musculus
6	45	71.4	525	11 Q8C1C1	Q8clc1 mus musculus
7	44	69.8	221	16 Q8DVI9	Q8dvi9 streptococc
8	42	66.7	512	13 Q7SY70	Q7sy70 xenopus lae
9	42	66.7	957	10 Q94GY5	Q94gy5 oryza sativ
10	41	65.1	465	1 Q9P9K8	Q9p9k8 methanosarc
11	41	65.1	480	17 Q8TRG4	Q8trg4 methanosarc
12	40	63.5	46	16 O07606	O07606 bacillus su
13	40	63.5	155	6 O77634	O77634 sus scrofa
14	40	63.5	181	16 Q7VFT8	Q7vft8 helicobacte
15	40	63.5	184	2 Q9ANV3	Q9anv3 lactobacill
16	40	63.5	328	10 Q8GYB6	Q8gyb6 arabidopsis

17	40	63.5	545	12 Q9WAT6	Q9wat6 chuzan viru
18	40	63.5	652	12 Q41929	Q41929 murid herpe
19	40	63.5	972	5 Q9XZ08	Q9xz08 drosophila
20	40	63.5	1356	10 Q9SH77	Q9sh77 arabidopsis
21	39	61.9	124	12 Q8JKN4	Q8jkn4 heliothis z
22	39	61.9	167	12 Q8VBP4	Q8vbp4 thogoto vir
23	39	61.9	167	12 Q8VBP5	Q8vbp5 thogoto vir
24	39	61.9	257	3 Q871L9	Q871l9 neurospora
25	39	61.9	320	12 Q86517	Q86517 rotavirus (
26	39	61.9	388	5 Q9N652	Q9n652 drosophila
27	39	61.9	388	5 Q9NGG6	Q9ngg6 drosophila
28	39	61.9	388	5 Q9N656	Q9n656 drosophila
29	39	61.9	388	5 Q9N6D3	Q9n6d3 drosophila
30	39	61.9	666	10 Q9LDM5	Q9ldm5 arabidopsis
31	39	61.9	753	16 Q8E6M8	Q8e6m8 streptococc
32	39	61.9	753	16 Q8E174	Q8e174 streptococc
33	39	61.9	790	5 Q9GSF3	Q9gsf3 podocoryne
34	39	61.9	808	4 Q96JN3	Q96jn3 homo sapien
35	39	61.9	1235	5 Q9VYL6	Q9vyl6 drosophila
36	39	61.9	1235	5 Q9GV19	Q9gv19 drosophila
37	39	61.9	1257	5 Q9GV18	Q9gv18 drosophila
38	39	61.9	1257	5 Q8IR79	Q8ir79 drosophila
39	38	60.3	117	5 Q9V8F5	Q9v8f5 drosophila
40	38	60.3	150	16 Q81X24	Q81x24 bacillus an
41	38	60.3	216	13 Q7SKM9	Q7skm9 brachydanio
42	38	60.3	221	4 Q9NP50	Q9np50 homo sapien
43	38	60.3	221	11 Q8C8M1	Q8c8m1 mus musculu
44	38	60.3	228	16 Q8RCE2	Q8rce2 thermonaer
45	38	60.3	277	11 P70361	P70361 mus musculu

ALIGNMENTS

RESULT 1

Q8UVZ4 ID Q8UVZ4 PRELIMINARY; PRT; 520 AA.

AC Q8UVZ4;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Lysosomal cofactor/neurotrophic factor prosaposin.

OS Brachydanio rerio (zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Danio.

OX NCBI_TaxID=7955;

RN [1]

RP SEQUENCE FROM N.A.

RA Seo H.-C., Lie O., Fjose A., O'Brien J.S., Kishimoto Y.;

RT "Cloning, expression and promoter analysis of zebrafish prosaposin.";

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF276996; AAL54381.1; -

DR GO; GO:0005764; C:lysosome; IEA.

DR GO; GO:0006665; P:sphingolipid metabolism; IEA.

DR InterPro; IPR003119; SapA.

DR InterPro; IPR007856; SapB_1.

DR InterPro; IPR008138; SapB_2.

DR InterPro; IPR008140; SapB_sub.

DR InterPro; IPR008373; Saposin.

DR InterPro; IPR008139; SaposinB.

DR Pfam; PF02199; SAPA; 2.

DR Pfam; PF05184; SapB_1; 3.

DR Pfam; PF03489; SapB_2; 4.

DR PRINTS; PR01797; SAPOSIN.

DR PRODOM; PD001732; SapB_sub; 3.

DR SMART; SM00162; SAPA; 2.

DR SMART; SM00118; SAPB; 4.

SQ SEQUENCE 520 AA; 57431 MW; F9E620F84BA41CB5 CRC64;

Query Match 79.4%; Score 50; DB 13; Length 520;
Best Local Similarity 63.6%; Pred. No. 0.62;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY      1 GDVQCDCIQMW 11
Db      186 GDVQCDCVTFI 196

RESULT 2
Q9DG82      PRELIMINARY;      PRT;      522 AA.
AC Q9DG82;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Prosaposin.
GN PSAP.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Altman N., Horowitz M.;
RT "The zebrafish prosaposin cDNA.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF108655; AAG32919.1; -.
DR ZFIN; ZDB-GENE-020108-1; psap.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SappB_1.
DR InterPro; IPR008138; SappB_2.
DR InterPro; IPR008140; SappB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SappB_1; 3.
DR Pfam; PF03489; SappB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SappB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPPB; 4.
SQ SEQUENCE 522 AA; 57671 MW; D3C15A305725C1CD CRC64;

Query Match      79.4%; Score 50; DB 13; Length 522;
Best Local Similarity 63.6%; Pred. No. 0.62;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 GDVQCDCIQMW 11
Db      186 GDVQCDCVTFI 196

RESULT 3
Q8BFQ1      PRELIMINARY;      PRT;      554 AA.
AC Q8BFQ1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Prosaposin.
GN PSAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NOD; TISSUE=Kidney, and Thymus;
RL MEDLINE=22354683; PubMed=12466851;
DR EMBL; AK009408; BAC25258.1; -.
DR MGD; MGI:1924193; 2310020A21Rik.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SappB_1.
DR InterPro; IPR008138; SappB_2.
DR InterPro; IPR008140; SappB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 1.
DR Pfam; PF05184; SappB_1; 3.
DR Pfam; PF03489; SappB_2; 3.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SappB_sub; 2.
DR SMART; SM00162; SAPA; 1.
DR SMART; SM00118; SAPPB; 3.
KW Hypothetical protein.
RP "Analysis of the mouse transcriptome based on functional annotation of

```

```

RT      60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK088369; BAC40308.1; -.
DR EMBL; AK089998; BAC41035.1; -.
DR MGD; MGI:197783; Psap.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SappB_1.
DR InterPro; IPR008138; SappB_2.
DR InterPro; IPR008140; SappB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SappB_1; 4.
DR Pfam; PF03489; SappB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SappB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPPB; 4.
SQ SEQUENCE 554 AA; 61050 MW; FF58DB79C7CC0C18 CRC64;

Query Match      74.6%; Score 47; DB 11; Length 554;
Best Local Similarity 70.0%; Pred. No. 2.3;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      2 DVCQDCIQMW 11
Db      195 DVCQDCMKLV 204

RESULT 4
Q8CINO      PRELIMINARY;      PRT;      402 AA.
AC Q8CINO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical saposin A-type domain/saposin type B containing
DE protein.
GN 2310020A21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Tongue;
RL MEDLINE=22354683; PubMed=12466851;
DR EMBL; AK009408; BAC25258.1; -.
DR MGD; MGI:1924193; 2310020A21Rik.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SappB_1.
DR InterPro; IPR008138; SappB_2.
DR InterPro; IPR008140; SappB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 1.
DR Pfam; PF05184; SappB_1; 3.
DR Pfam; PF03489; SappB_2; 3.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SappB_sub; 2.
DR SMART; SM00162; SAPA; 1.
DR SMART; SM00118; SAPPB; 3.
KW Hypothetical protein.
RP "Analysis of the mouse transcriptome based on functional annotation of

```

```

Query Match      71.4%; Score 45; DB 11; Length 402;
Best Local Similarity 54.5%; Pred. No. 4;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GDVQCDCIQMV 11
DB 64 GAVCHDCVQLI 74

RESULT 5
Q8BUJ5 PRELIMINARY; PRT; 449 AA.
AC Q8BUJ5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical saposin A-type domain/saposin type B containing protein.
GN 2310020A21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eye;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
RT Nature 420:563-573(2002).
RL Nature 420:563-573(2002).
DR EMBL; AK078699; BAC37363.1; -.
DR GO; GO:0005764; C:lysosome; IEA.
DR MGD; MGI:1924193; 2310020A21RIK.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR GO; GO:005764; C:lysosome; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SapB_1; 4.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 4.
KW Hypothetical protein.
SQ SEQUENCE 449 AA; 48705 MW; 377CC42A475B292F CRC64;

Query Match      71.4%; Score 45; DB 11; Length 449;
Best Local Similarity 54.5%; Pred. No. 4.4;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GDVQCDCIQMV 11
DB 187 GAVCHDCVQLI 197

RESULT 6
Q8C1C1 PRELIMINARY; PRT; 525 AA.
AC Q8C1C1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical saposin A-type domain/saposin type B containing protein.
GN 2310020A21RIK.
OS Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
RT Nature 420:563-573(2002).
RL Nature 420:563-573(2002).
DR EMBL; AK028455; BAC25961.1; -.
DR MGD; MGI:1924193; 2310020A21RIK.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SapB_1; 4.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 4.
KW Hypothetical protein.
SQ SEQUENCE 525 AA; 57350 MW; 374F6050CDC4D223 CRC64;

Query Match      71.4%; Score 45; DB 11; Length 525;
Best Local Similarity 54.5%; Pred. No. 5;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GDVQCDCIQMV 11
DB 187 GAVCHDCVQLI 197

RESULT 7
Q8DV19 PRELIMINARY; PRT; 221 AA.
AC Q8DV19;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative late competence protein.
DR SMU.499.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL; AE014895; AAN58244.1; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000345; CytC heme BS.
DR PROSITE; PS00190; CYTOCHROME_C; 2.
KW Complete proteome.
SQ SEQUENCE 221 AA; 26017 MW; C3C68D1C6A7F697D CRC64;

Query Match      69.8%; Score 44; DB 16; Length 221;

```

```

Best Local Similarity 85.7%; Pred. No. 3.5;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDVQCDC 7
Db 54 GDICQDC 60

RESULT 8
Q7SY70
AC Q7SY70 PRELIMINARY; PRT; 512 AA.
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T.I., Max S.I., Wang J.J., Haie H.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RA Klein S., Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC054988; AAI54988.1; -
KW Hypothetical protein.
SQ SEQUENCE 512 AA; 57555 MW; 57CFA5E2093F6FB1 CRC64;

Query Match 66.7%; Score 42; DB 13; Length 512;
Best Local Similarity 45.5%; Pred. No. 17;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 GDVQCDCQI 11
Db 192 GDICNDCTTLLI 202

RESULT 9
Q94GY5
AC Q94GY5 PRELIMINARY; PRT; 957 AA.
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Teitrit T.,
RA Riggs F., Heiao J., Zismann V., Blunt S., Pal G., VanAken S.E.,
RA Unterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,
RA White O., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNBA0018H01 genomic sequence.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RX EMBL; AC087181; AAK38494.1; -
DR Gramene; Q94GY5; -
DR InterPro; IPR007656; DUF593.
DR Pfam; PF04576; DUF593; 1.
KW Hypothetical protein.
SQ SEQUENCE 957 AA; 105813 MW; 6C569D73B33C3C38 CRC64;

Query Match 66.7%; Score 42; DB 10; Length 957;
Best Local Similarity 62.5%; Pred. No. 30;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDVQCDCI 8
Db 113 GDWCEDCV 120

RESULT 10
Q9P9K8
AC Q9P9K8 PRELIMINARY; PRT; 465 AA.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative vanadium dinitrogenase alpha subunit.
GN VNF0.
OS Methanosarcina barkeri.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2208;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43241 / DSM 1538 / 227;
RX MEDLINE=20270156; PubMed=10809706;
RA Chien Y.T., Auerbuch V., Brabban A.D., Zinder S.H.;
RT "Analysis of genes encoding an alternative nitrogenase in the archaeon
RT Methanosarcina barkeri 227.";
RL J. Bacteriol. 182:3247-3253(2000).
DR EMBL; AF254784; AAF72180.1; -
DR HSSP; P00467; 1MIO.
DR GO; GO:0016734; F:molybdenum-iron nitrogenase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0009339; P:nitrogen fixation; IEA.
DR InterPro; IPR000318; Nitrogenase compl.
DR InterPro; IPR005974; Nitrogenase.
DR InterPro; IPR00510; Oxidized nitro; 1.
DR TIGRfams; TIGR01284; alt nitrog alph; 1.
DR PROSITE; PS00699; NITROGENASE_1; 1.
SQ SEQUENCE 465 AA; 52841 MW; B9312F97C876D9A3 CRC64;

```


Query Match 65.1%; Score 41; DB 1; Length 465;
 Best Local Similarity 63.6%; Pred. No. 24;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11
 | | | | |
 Db 59 GGVVKDCIQMI 69

RESULT 11

Q8TRG4 PRELIMINARY; PRT; 480 AA.
 AC Q8TRG4;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Nitrogenase, subunit alpha.
 GN VNF0 OR M1216.
 OS Methanosaerina acetivorans.
 OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
 OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=11932238;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA FitzHugh W., Calvo S., Engels-Thomann N., Dearellano K., Johnson R.,
 RA Allen N., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
 RA Linton L., McGowan P., Ranganathan K., Ralston J., Smith K.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Graham D.A., Guss A.M.,
 RA Hederich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;
 RA "The genome of Methanosarcina acetivorans reveals extensive metabolic
 RT and physiological diversity";
 RL Genome Res. 12:532-542(2002).
 DR EMBL; AE010789; AAM04635.1; --
 DR HSSP; P07328; 1WIN.
 DR GO; GO:0016734; F:molybdenum-iron nitrogenase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0009399; P:nitrogen fixation; IEA.
 DR InterPro; IPR000318; Nitrogenase compl.
 DR InterPro; IPR005974; Nitrogenase.
 DR InterPro; IPR000510; Oxidized nitrogn.
 DR Pfam; PF00148; Oxidized nitro; 1.
 DR TIGRFAMs; TIGR01284; alt nitroq alph; 1.
 DR PROSITE; PS00699; NITROGENASE_1; 1.
 KW Complete proteome.
 SQ SEQUENCE 480 AA; 54598 MW; E8D57106E82474D CRC64;

Query Match 65.1%; Score 41; DB 17; Length 480;
 Best Local Similarity 63.6%; Pred. No. 24;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11
 | | | | |
 Db 59 GGVVKDCIQMI 69

RESULT 12

O07606 PRELIMINARY; PRT; 46 AA.
 AC O07606;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein ynfH.
 GN YHFH.

OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Noback M.A., Terpstra P., Holsappel S., Venema G., Bron S.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Haeck J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kaahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Srokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RL subtilis";
 RL Nature 390:249-256(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y14083; CAA74529.1; --
 DR EMBL; Z99109; CAB12863.1; --
 DR PIR; E69830; E69830.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 46 AA; 5294 MW; AF80B4AA71C18061 CRC64;

Query Match 63.5%; Score 40; DB 16; Length 46;
 Best Local Similarity 50.0%; Pred. No. 4.6;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQM 10
 | | | | |
 Db 34 GNICNDICV 43

RESULT 13

O77634 PRELIMINARY; PRT; 155 AA.
 ID O77634
 AC O77634;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cellular disintegrin (Fragment).
 GN ADAM-9.
 OS Sus scrofa (Pig).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99357011; PubMed=10429942;
RA Flannery C.R., Little C.B., Caterson B., Hughes C.E.;
RT "Effects of culture conditions and exposure to catabolic stimulators
RT (IL-1 and retinoic acid) on the expression of matrix
RT metalloproteinases (MMPs) and disintegrin metalloproteinases (ADAMs)
RT by articular cartilage chondrocytes.";
RL Matrix Biol. 18:225-237(1999).
DR EMBL; AF069646; AAC23530.1; -.
DR HSSP; P18619; 1FVL.
DR GO; GO:0007229; P: integrin-mediated signaling pathway; IEA.
DR InterPro; IPR006586; ADAM cysteine.
DR InterPro; IPR001762; Disintegrin.
DR Pfam; PF00200; disintegrin; 1.
DR PRINTS; PD00289; disintegrin; 1.
DR ProDom; PD000664; Disintegrin; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR PROSITE; PS02114; DISINTEGRIN_2; 1.
KW Integrin.
FT NON_TER 1
FT NON_TER 155
FT NON_TER 155
SQ SEQUENCE 155 AA; 16890 MW; 48A7CA6075F9390E CRC64;

Query Match 63.5%; Score 40; DB 6; Length 155;
Best Local Similarity 85.7%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GDVQCQC 7
DB 19 GDCCQDC 25

RESULT 14
QVFT8
ID Q7VFT8 PRELIMINARY; PRT; 181 AA.
AC Q7VFT8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN HH1587.
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 51449 / 3B1;
RX MEDLINE=22709201; PubMed=12810954;
RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,
RA Bell M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
DR EMBL; AE017148; AAP78184.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 181 AA; 21202 MW; 4212C4E40E6CD486 CRC64;

Query Match 63.5%; Score 40; DB 16; Length 181;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GDVQCQCIC 9
DB 75 GDLIQDCVQ 83

RESULT 15
Q9ANV3
ID Q9ANV3 PRELIMINARY; PRT; 184 AA.
AC Q9ANV3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Putative ComF3 protein.
GN COMF3.
OS Lactobacillus delbrueckii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1584;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC11842;
RA van de Guchte M., Dervyn R., Ehrlich S.D., Maguin E.;
RT "L. bulgaricus ymdA - rf2 region.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF320250; AAK00332.1; -.
SQ SEQUENCE 184 AA; 20908 MW; 6ECF46DAEC33C881 CRC64;

Query Match 63.5%; Score 40; DB 2; Length 184;
Best Local Similarity 71.4%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDVQCQC 7
DB 57 GEICQDC 63

Search completed: May 5, 2004, 13:33:42
Job time : 2.09687 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:16:48 ; Search time 0.271605 Seconds
(without alignments)
2108.841 Million cell updates/sec

Title: US-09-743-684A-19

Perfect score: 63
Sequence: 1 GDVQCDCIQW 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	100.0	80	1 SAP_PIG	P81405 sus scrofa
2	63	100.0	524	1 SAP_HUMAN	P07602 h proactiva
3	55	87.3	525	1 SAP_BOVIN	P26779 b proactiva
4	50	79.4	518	1 SAP_CHICK	O13035 gallus gall
5	47	74.6	554	1 SAP_RAT	P10960 rattus norv
6	47	74.6	557	1 SAP_MOUSE	Q61207 mus musculu
7	39	61.9	512	1 VENV_THOXY	P28977 thogoto vlr
8	39	61.9	1696	1 PCRS_BRACL	Q94315 branchiosteo
9	38	60.3	355	1 CMG2_SCHPO	P32434 schizosacch
10	37	58.7	74	1 Y777_TREPA	Q83756 treponema p
11	37	58.7	374	1 NADA_HALNI	Q9hnu1 halobacteri
12	37	58.7	422	1 U183_HUMAN	Q9bsu1 homo sapien
13	37	58.7	422	1 U183_MOUSE	Q922r1 mus musculu
14	37	58.7	422	1 U183_RAT	Q08654 rattus norv
15	37	58.7	748	1 CLPE_LACLA	Q9c109 lactococcus
16	37	58.7	748	1 CLPE_LACLA	Q9s522 lactococcus
17	37	58.7	1015	1 ITA4_TROME	Q94724 drosophila
18	37	58.7	5147	1 FAT_DROME	P33450 drosophila
19	36	57.1	194	1 BCIB_HUMAN	Q9hnd3 homo sapien
20	36	57.1	330	1 EMB_MOUSE	P21995 mus musculu
21	36	57.1	425	1 CND0_HUMAN	Q84806 homo sapien
22	36	57.1	425	1 CND0_MOUSE	Q8hnu4 mus musculu
23	36	57.1	511	1 GINB_PSEPL	P18126 pseudomonas
24	36	57.1	564	1 TM16_HUMAN	Q95361 homo sapien
25	36	57.1	757	1 COMP_HUMAN	P49747 homo sapien
26	36	57.1	819	1 AD09_HUMAN	Q14443 homo sapien
27	35.5	56.3	1416	1 YN81_CAEEL	Q03610 caenorhabdi
28	35	55.6	48	1 R332_MYCGE	Q9z882 mycoplasma
29	35	55.6	48	1 R332_MYCPV	P56850 mycoplasma
30	35	55.6	103	1 THB1_THORP	Q06884 rhodnius pr
31	35	55.6	175	1 CEN_ARATH	Q922v5 arabidopsis
32	35	55.6	181	1 Y4AS_RHISN	P55366 rhizobium s
33	35	55.6	199	1 AA27_MOUSE	P56873 mus musculu

34	35	55.6	228	1 VIF_OMVVS	P16902 ovine lenti
35	35	55.6	230	1 VIF_VILV	P03403 vlna lenti
36	35	55.6	230	1 VIF_VILV1	P23430 vlna lenti
37	35	55.6	285	1 T13E_HUMAN	Q94275 homo sapien
38	35	55.6	309	1 T13B_MOUSE	Q9wu72 mus musculu
39	35	55.6	477	1 TM17_MOUSE	Q7cpm3 mus musculu
40	35	55.6	477	1 TM17_RAT	Q9wv59 rattus norv
41	35	55.6	498	1 YAC3_SCHPO	Q09817 schizosacch
42	35	55.6	752	1 CLPE_STRPN	P35594 streptococc
43	35	55.6	867	1 ARX7_MOUSE	Q84411 mus musculu
44	35	55.6	892	1 ARX7_HUMAN	O15265 homo sapien
45	35	55.6	961	1 TSP4_HUMAN	P35443 homo sapien

ALIGNMENTS

RESULT 1

ID	SAP_PIG	STANDARD	PRT	80 AA.
AC	P81405			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Sapospin B (Cerebroside sulfate activator) (CS-ACT) (Non-specific activator) (Sphingolipid activator protein 1) (SAP-1).			
OS	Sus scrofa (pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OX	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
NC	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE OF 1-79.			
RC	TISSUE=Kidney;			
RX	MEDLINE=93229506; PubMed=8471613;			
RA	Stevens R.L., Fauli K.F., Conklin K.A., Green B.N., Fluharty A.L.;			
RT	"porcine cerebroside sulfate activator: further structural characterization and disulfide identification.";			
RL	Biochemistry 32:4051-4059(1993).			
RN	[2]			
RP	SEQUENCE OF 1-64.			
RC	TISSUE=Kidney;			
RX	MEDLINE=92222651; PubMed=1562358;			
RA	Fluharty A.L., Katona Z., Meek W.B., Frei K., Fowler A.V.;			
RT	"The cerebroside sulfate activator from pig kidney: purification and molecular structure.";			
RL	Biochem. Med. Metab. Biol. 47:66-85(1992).			
RN	[3]			
RP	STRUCTURE OF CARBOHYDRATE ON ASN-21.			
RX	MEDLINE=21110404; PubMed=1180632;			
RA	Fauli K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,			
RT	Stevens R.L., Fluharty C.B., Fluharty A.L.;			
RL	"Structure of the asparagine-linked sugar chains of porcine kidney and human urine cerebroside sulfate activator protein.";			
RN	J. Mass Spectrom. 35:1416-1424(2000).			
RP	[4]			
RC	TISSUE=Kidney;			
RX	MEDLINE=99441404; PubMed=10510427;			
RA	Fauli K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,			
RT	Krutchinsky A.N., Standing K.G., Waring A.J., Stevens R.L.,			
RL	Fluharty C.B., Fluharty A.L.;			
RT	"Cerebroside sulfate activator protein (Sapospin B): chromatographic and electrospray mass spectrometric properties.";			
RL	J. Mass Spectrom. 34:1040-1054(1999).			
CC	-1- FUNCTION: Sapospin B stimulates the hydrolysis of galacto-			
CC	cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GMI			
CC	gangliosides by beta-galactosidase (EC 3.2.1.23) and			
CC	globosialosylceramide by alpha-galactosidase A (EC 3.2.1.22).			
CC	Sapospin B forms a solubilizing complex with the substrates of the			
CC	sphingolipid hydrolases.			
CC	-1- SUBUNIT: Sapospin B is a homodimer (by similarity).			
CC	-1- PTM: The one residue extended Sapospin B-Val is only found in a			
CC	minority of the chains.			

CC -1- SIMILARITY: Contains 1 saposin B-type domain.
 DR GlycosultideB; P81405; -;
 DR InterPro: IPR007856; Saps_1.
 DR InterPro: IPR008138; Saps_2.
 DR InterPro: IPR008373; Saposin.
 DR InterPro: IPR008139; SaposinB.
 DR Pfam: PF05184; SapsB_1; 1.
 DR Pfam: PF03489; SapsB_2; 1.
 DR PRINTS: PR01797; SAPOSIN.
 DR SMART: SM00118; SapsB; 1.
 DR Glycoprotein; sphingolipid metabolism.
 FT CHAIN 1 79 SAPOSIN B-VAL.
 FT DOMAIN 1 80 SAPOSIN-LIKE TYPE B.
 FT DISULFID 4 77
 FT DISULFID 7 71
 FT DISULFID 36 47
 FT CARBOHYD 21 21
 FT N-LINKED (GLCNAC. . .) (COMPLEX).
 FT /FTID=CAR_000177.
 FT EF7BA249B63E789C CRC64;
 SQ SEQUENCE 80 AA; 8949 MW; /FTID=CAR_000177.
 Query Match 100.0%; Score 63; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 0.00038;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GDVQCDCIQMV 11
 DB 1 GDVQCDCIQMV 11
 RESULT 2
 SLP_HUMAN
 ID_SAP_HUMAN STANDARD; PRT; 524 AA.
 AC P07602; P07292; P15793; P78538; P78541; P78546; P78547; P78558;
 AC Q92739; Q92740; Q92741; Q92742;
 DT 01-APR-1998 (Rel. 07. Created)
 DT 01-APR-1998 (Rel. 14. Last sequence update)
 DT 10-OCT-2003 (Rel. 42. Last annotation update)
 DE Proactivator polypeptide precursor [Contains: Saposin A (Protein A);
 DE Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside
 DE sulfate activator) (CSAct) (Dispersin) (Sulfatide/GM1 activator);
 DE Saposin C (Co-beta-glucosidase) (A1 activator) (Glucosylceramidase
 DE activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D
 DE (Protein C) (Component C)].
 GN PSAP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=90129043; PubMed=2515150;
 RA Roman E.G., Grabowski G.A.;
 RT "Molecular cloning of a human co-beta-glucosidase cDNA: evidence that
 RT four sphingolipid hydrolase activator proteins are encoded by single
 RT genes in humans and rats.";
 RL Genomics 5:486-492(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89255151; PubMed=2498298;
 RA Nakano T., Sandhoff K., Suemmer J., Christomanou H., Suzuki K.;
 RT "Structure of full-length cDNA coding for sulfatide activator, a
 RT co-beta-glucosidase and two other homologous proteins: two alternate
 RT forms of the sulfatide activator.";
 RL J. Biochem. 105:152-154(1989).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM SAP-WU-0).
 RC TISSUE=Brain, Eye, and Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rahn S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
 RA Boeak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Wuzny D.W., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Faney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherzer A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 59-125 AND 304-513 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=91192146; PubMed=2013321;
 RA Holtzman H., Sandhoff K., Fuerst W., Kwon H.Y., Schnabel D.,
 RA Suzuki K.;
 RT "The organization of the gene for the human cerebroside sulfate
 RT activator protein.";
 RL FEBS Lett. 280:267-270(1991).
 RN [5]
 RP SEQUENCE OF 164-524 FROM N.A.
 RX MEDLINE=88068647; PubMed=2825202;
 RA Dewi N.N., Wenger D.A., O'Brien J.S.;
 RT "Nucleotide sequence of cloned cDNA for human sphingolipid activator
 RT protein 1 precursor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8652-8656(1987).
 RN [6]
 RP PARTIAL SEQUENCE OF 60-142.
 RX MEDLINE=89240739; PubMed=2717620;
 RA Morimoto S., Martin B.M., Yamamoto Y., Kretz K.A., O'Brien J.S.,
 RA Kishimoto Y.;
 RT "Saposin A: second cerebroside activator protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3389-3393(1989).
 RN [7]
 RP SEQUENCE OF 195-263 FROM N.A.
 RX MEDLINE=86130593; PubMed=2868718;
 RA Dewi N.N., Wenger D.A., Fujibayashi S., Donoviel M., Esch F.,
 RA Hill F., O'Brien J.S.;
 RT "Molecular cloning of the sphingolipid activator protein-1 (SAP-1),
 RT the sulfate sulfatase activator.";
 RL Biochem. Biophys. Res. Commun. 134:989-994(1986).
 RN [8]
 RP SEQUENCE OF 195-274.
 RC TISSUE=Kidney;
 RX MEDLINE=91006165; PubMed=2209618;
 RA First W., Schubert J., Machleidt W., Meyer H.E., Sandhoff K.;
 RT "The complete amino-acid sequences of human ganglioside GM2 activator
 RT protein and cerebroside sulfate activator protein.";
 RL Eur. J. Biochem. 192:709-714(1990).
 RN [9]
 RP SEQUENCE OF 195-274.
 RX MEDLINE=89207118; PubMed=3242555;
 RA Kleinschmidt T., Christomanou H., Braunitzer G.;
 RT "Complete amino-acid sequence of the naturally occurring A2 activator
 RT protein for enzymic sphingomyelin degradation: identity to the
 RT sulfatide activator protein (SAP-1).";
 RL Biol. Chem. Hoppe-Seyler 369:1361-1365(1988).
 RN [10]
 RP SEQUENCE OF 311-390.
 RX MEDLINE=88163077; PubMed=3442600;
 RA Kleinschmidt T., Christomanou H., Braunitzer G.;
 RT "Complete amino-acid sequence and carbohydrate content of the
 RT naturally occurring glucosylceramide activator protein (A1 activator)
 RT absent from a new human Gaucher disease variant.";
 RL Biol. Chem. Hoppe-Seyler 368:1571-1578(1987).
 RN [11]

RP SEQUENCE OF 407-484.
 RX MEDLINE=89000190; PubMed=3048308;
 RA Furst W., Machleidt W., Sandhoff K.;
 RT "The precursor of sulfatide activator protein is processed to three
 RT different proteins.";
 RL Biol. Chem. Hoppe-Seyler 369:317-328(1988).
 RN [12]
 RP PARTIAL SEQUENCE OF 405-484.
 RX MEDLINE=89025876; PubMed=2845979;
 RA Morimoto S., Martin B.M., Kishimoto Y., O'Brien J.S.;
 RT "Saposin D: a sphingomyelinase activator.";
 RL Biochem. Biophys. Res. Commun. 156:403-410(1988).
 RN [13]
 RP SEQUENCE OF 17-26.
 RC TISSUE=Milk;
 RX MEDLINE=92068206; PubMed=1958198;
 RA Kondoh K., Haheno T., Sano A., Kakimoto Y.;
 RT "Isolation and characterization of prosaposin from human milk.";
 RL Biochem. Biophys. Res. Commun. 181:286-292(1991).
 RN [14]
 RP PARTIAL SEQUENCE (SAPOSIN B), AND STRUCTURE OF CARBOHYDRATES.
 RC TISSUE=Urine;
 RX MEDLINE=20032116; PubMed=10562467;
 RA Fluharty A.L., Lombardo C., Louis A., Stevens R.L., Whitelegge J.P.,
 RA Waring A.J., To T., Fluharty C.B., Faull K.F.;
 RT "Preparation of the cerebroside sulfate activator (CSact or saposin B)
 RT from human urine.";
 RL Mol. Genet. Metab. 68:391-403(1999).
 RN [15]
 RP STRUCTURE OF CARBOHYDRATE ON ASN-215.
 RX MEDLINE=21110404; PubMed=1180632;
 RA Faull K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,
 RA Stevens R.L., Fluharty C.B., Fluharty A.L.;
 RT "Structure of the asparagine-linked sugar chains of porcine kidney and
 RT human urine cerebroside sulfate activator protein.";
 RL J. Mass Spectrom. 35:1416-1424(2000).
 RN [16]
 RP SAPOSIN D DISULFIDE BONDS.
 RX MEDLINE=99337688; PubMed=10406958;
 RA Tatci M., Salvio R., Clafioni F., Pucci P., Andolfo A.,
 RA Amoresano A., Vaccaro A.M.;
 RT "Structural and membrane-binding properties of saposin D.";
 RL Eur. J. Biochem. 263:486-494(1999).
 RN [17]
 RP SAPOSIN B DISULFIDE BONDS.
 RX MEDLINE=22398398; PubMed=12510003;
 RA Ahn V.E., Faull K.F., Whitelegge J.P., Higginson J., Fluharty A.L.,
 RA Prive G.G.;
 RT "Expression, purification, crystallization, and preliminary X-ray
 RT analysis of recombinant human saposin B.";
 RL Protein Expr. Purif. 27:186-193(2003).
 RN [18]
 RP MASS SPECTROMETRY.
 RC TISSUE=Urine;
 RX MEDLINE=99441404; PubMed=10510427;
 RA Faull K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,
 RA Krutcheny A.N., Standing K.G., Waring A.J., Stevens R.L.,
 RA Fluharty C.B., Fluharty A.L.;
 RT "Cerebroside sulfate activator protein (Saposin B): chromatographic
 RT and electrospray mass spectrometric properties.";
 RL J. Mass Spectrom. 34:1040-1054(1999).
 RN [19]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 195-273, AND MUTAGENESIS OF
 RP ILE-240.
 RX MEDLINE=22406333; PubMed=12518053;
 RA Ahn V.E., Faull K.F., Whitelegge J.P., Fluharty A.L., Prive G.G.;
 RT "Crystal structure of saposin B reveals a dimeric shell for lipid
 RT binding.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:38-43(2003).
 RN [20]
 RP REVIEW ON MLD VARIANTS.
 RX MEDLINE=95170731; PubMed=7866401;
 RA Gieselmann V., Zlotogora J., Harris A., Wenger D.A., Morris C.P.;

RT "Molecular genetics of metachromatic leukodystrophy.";
 RL Hum. Mutat. 4:233-242(1994).
 RN [21]
 RP VARIANT MLD ILE-217.
 RX MEDLINE=90147748; PubMed=2302219;
 RA Rafi M.A., Zhang X.-L., Degala G., Wenger D.A.;
 RT "Detection of a point mutation in sphingolipid activator protein-1
 RT mRNA in patients with a variant form of metachromatic
 RT leukodystrophy.";
 RL Biochem. Biophys. Res. Commun. 166:1017-1023(1990).
 RN [22]
 RP SEQUENCE FROM N.A., AND VARIANT MLD ILE-217.
 RX MEDLINE=90207231; PubMed=2320574;
 RN [23]
 Qy Query Match 100.0%; Score 63; DB 1; Length 524;
 Best Local Similarity 100.0%; Pred. No. 0.0023;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 195 GDVCDCTQWV 11
 1 GDVCDCTQWV 11
 195 GDVCDCTQWV 205
 RESULT 3
 SAP_BOVIN STANDARD; PRT; 525 AA.
 AC P26779; Q9N2G4;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Proactivator polypeptide precursor [contains: Saposin A (Protein A);
 DE Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside
 DE sulfate activator) (CSact) (Dispersin) (Sulfatide/GM1 activator);
 DE Saposin C (Co-beta-glucosidase) (AI activator) (Glucosylceramidase
 DE activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D
 DE (Protein C) (Component C)].
 GN PSAP.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANTS.
 RC TISSUE=Mammary gland;
 RA Azuma N., Yoshida K.;
 RT "RT-PCR cloning of bovine prosaposin.";
 RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 312-391.
 RC TISSUE=Spleen;
 RX MEDLINE=92207994; PubMed=1554743;
 RA Sano A., Mizuno T., Kondoh K., Haheno T., Ueno S.-I., Kakimoto Y.,
 RA Morita N.;
 RT "Saposin-C from bovine spleen: complete amino acid sequence and
 RT relation between the structure and its biological activity.";
 RL Biochim. Biophys. Acta 1120:75-80(1992).
 RN [3]
 RP FUNCTION: The lysosomal degradation of sphingolipids takes place
 CC by the sequential action of specific hydrolases. Some of these
 CC enzymes require specific low-molecular mass, non-enzymic proteins:
 CC the sphingolipids activator proteins (coproteins) (By similarity).
 CC
 CC -!- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
 CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
 CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
 CC Saposin C apparently acts by combining with the enzyme and acidic
 CC lipid to form an activated complex, rather than by solubilizing
 CC the substrate.
 CC
 CC -!- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
 CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1
 CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
 CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
 CC Saposin B forms a solubilizing complex with the substrates of the
 CC sphingolipid hydrolases (By similarity).

```

CC -1- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
CC activator (EC 3.1.4.12) (By similarity).
CC -1- SUBUNIT: Saposin B is a homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Lysosomal.
CC -1- PTM: This precursor is proteolytically processed to 4 small
CC peptides, which are similar to each other and are sphingolipid
CC hydrolase activator proteins (By similarity).
CC -1- SIMILARITY: Contains 2 saposin A-type domains.
CC -1- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB036791; BAA95677.1; -.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; Sapa_1.
DR InterPro; IPR008138; Sapa_2.
DR InterPro; IPR008140; Sapa_sub.
DR InterPro; IPR008373; Saposinb.
DR Pfam; PF02199; Sapa_1; 2.
DR Pfam; PF05184; Sapa_1; 4.
DR Pfam; PF03489; Sapa_2; 4.
DR PRINTS; PR01797; Saposin.
DR ProDom; PD001732; Sapa_sub; 3.
DR SMART; SM00162; Sapa_2.
DR SMART; SM00118; Sapa_4.
DR Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat.
KW SIGNAL; 16
FT PROPEP 1 58
FT CHAIN 60 142 SAPOSIN A.
FT PROPEP 144 195
FT CHAIN 196 275 SAPOSIN B.
FT PROPEP 277 310
FT CHAIN 312 392 SAPOSIN C.
FT PROPEP 393 404
FT CHAIN 406 487
FT PROPEP 489 525 SAPOSIN D.
FT DOMAIN 21 54
FT DOMAIN 59 142 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 194 276 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 312 393 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 406 487 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 492 525 SAPOSIN-LIKE TYPE A 2.
FT DISULFD 63 138
FT DISULFD 66 132
FT DISULFD 66 132 BY SIMILARITY.
FT DISULFD 94 106 BY SIMILARITY.
FT DISULFD 199 272 BY SIMILARITY.
FT DISULFD 202 242 BY SIMILARITY.
FT DISULFD 231 262 BY SIMILARITY.
FT DISULFD 316 389 BY SIMILARITY.
FT DISULFD 319 383 BY SIMILARITY.
FT DISULFD 347 358 BY SIMILARITY.
FT DISULFD 410 483 BY SIMILARITY.
FT DISULFD 413 477 BY SIMILARITY.
FT DISULFD 441 452 BY SIMILARITY.
FT CARBOHYD 80 80
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 427 427 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 127 127 H -> R.
FT VARIANT 260 263 MQPK -> IRR.
FT CONFLICT 317 317 E -> Q (IN REF. 2).
FT CONFLICT 367 367 R -> S (IN REF. 1).
SQ SEQUENCE 525 AA; 58120 MW; 293AFC0FB9C4FA99 CRC64;
Query Match 87.3%; Score 55; DB 1; Length 525;

```

```

Best Local Similarity 81.8%; Pred. No. 0.05;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 GDVCDCTQMV 11
Db 196 GNVCDCTQLV 206
RESULT 4
SAP_CHICK STANDARD; PRT; 518 AA.
ID SAP_CHICK
AC O13035.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Proctivator polypeptide precursor [Contains: Saposin A; Saposin B;
DE Saposin C; Saposin D].
GN PSAP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_Taxid=9031;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 194-203.
RC TISSUE=Brain, and Liver;
RX MEDLINE=98129745; PubMed=9461526;
RA Azuma N., Seo H.-C., Lie O., Fu O., Gould R.M., Hiraiwa M., Burt D.W.,
RA Paton I.R., Morrice D.R., O'Brien J.S., Kishimoto Y.;
RT "Cloning, expression and map assignment of chicken prosaposin.";
RL Biochem. J. 330:321-327 (1998).
RP SEQUENCE FROM N.A.
RA Altman N., Horowitz M.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: The lysosomal degradation of sphingolipids takes place
CC by the sequential action of specific hydrolases. Some of these
CC enzymes require specific low-molecular mass, non-enzymic proteins:
CC the sphingolipids activator proteins (coproteins) (By similarity).
CC -1- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
CC Saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate (By similarity).
CC -1- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1
CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
CC Saposin B forms a solubilizing complex with the substrates of the
CC sphingolipid hydrolases (By similarity).
CC -1- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
CC activator (EC 3.1.4.12) (By similarity).
CC -1- SUBUNIT: Saposin B is a homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Lysosomal (By similarity).
CC -1- PTM: This precursor is proteolytically processed to 4 small
CC peptides, which are similar to each other and are sphingolipid
CC hydrolase activator proteins (By similarity).
CC -1- SIMILARITY: Contains 2 saposin A-type domains.
CC -1- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB036791; BAA19914.1; -.
DR EMBL; AF108656; AAF05899.1; -.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; Sapa_1.

```

```

DR InterPro; IPR008138; SApB_2.
DR InterPro; IPR008140; SApB sub.
DR InterPro; IPR008373; SApOSin.
DR InterPro; IPR008139; SApOSinB.
DR Pfam; PF02199; SApA_2.
DR Pfam; PF05184; SApB_1; 4.
DR Pfam; PF03489; SApB_2; 4.
DR PRINTS; PR01797; SApOSin.
DR ProDom; PD001732; SApB sub; 3.
KM Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat;
FT SIGNAL 1 17 POTENTIAL.
FT PROPEP 18 60
FT CHAIN 61 143 SApOSIN A.
FT PROPEP 145 193
FT CHAIN 194 276 SApOSIN B.
FT PROPEP 278 305 SApOSIN C.
FT CHAIN 307 387
FT PROPEP 389 398 SApOSIN D.
FT CHAIN 399 480
FT PROPEP 482 518
FT DOMAIN 22 55 SApOSIN-LIKE TYPE A 1.
FT DOMAIN 60 143 SApOSIN-LIKE TYPE B 1.
FT DOMAIN 193 277 SApOSIN-LIKE TYPE B 2.
FT DOMAIN 307 388 SApOSIN-LIKE TYPE B 3.
FT DOMAIN 399 480 SApOSIN-LIKE TYPE B 4.
FT DOMAIN 485 518 SApOSIN-LIKE TYPE A 2.
FT DISULFID 64 139 BY SIMILARITY.
FT DISULFID 67 133 BY SIMILARITY.
FT DISULFID 95 107 BY SIMILARITY.
FT DISULFID 197 273 BY SIMILARITY.
FT DISULFID 200 267 BY SIMILARITY.
FT DISULFID 229 340 BY SIMILARITY.
FT DISULFID 311 384 BY SIMILARITY.
FT DISULFID 314 378 BY SIMILARITY.
FT DISULFID 342 353 BY SIMILARITY.
FT DISULFID 403 476 BY SIMILARITY.
FT DISULFID 406 470 BY SIMILARITY.
FT DISULFID 434 445 BY SIMILARITY.
FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 94 94 R -> T (IN REF. 2).
FT CONFLICT 486 486 E -> D (IN REF. 2).
SQ SEQUENCE 518 AA; 57601 MW; B803000B91C3963 CRC64;

Query Match Best Local Similarity 79.4%; Score 50; DB 1; length 518;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 DVCODCIOMV 11
DB 195 DVCODCIRLV 204

RESULT 5
SAP_RAT STANDARD; PRT; 554 AA.
AC P10960; Q62841; Q64190;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN PSAP OR SGP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCB1_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Sertoli cells;
RX MEDLINE=89000647; PubMed=3048385;

```

```

RA Collard M.W., Sylvestre S.R., Tsuruta J.K., Griswold M.D.;
RT "Biosynthesis and molecular cloning of sulfated glycoprotein 1
RT secreted by rat Sertoli cells: sequence similarity with the
RT 70-kilodalton precursor to sulfatide/GM1 activator.";
RL Biochemistry 27:4557-4564(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96128541; PubMed=8573994;
RA Morales C.R., El-Alfy M., Zhao Q., Igdoura S.A.;
RT "Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in
RT Sertoli cells.";
RL Histol. Histopathol. 10:1023-1034(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RX MEDLINE=96175245; PubMed=8601692;
RA Morales C.R., El-Alfy M., Zhao Q., Igdoura S.A.;
RT "Expression and tissue distribution of rat sulfated glycoprotein-1
RT (prosaposin) ";
RL J. Histochem. Cytochem. 44:327-337(1996).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: Contains 2 saposin A-type domains.
CC -1- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M19936; AAA42136.1; -
CC EMBL; S81353; AAB36042.2; -
CC EMBL; S81373; AAB36233.2; -
CC PIR; A28716; A28716.
DR InterPro; IPR003119; SApA.
DR InterPro; IPR007856; SApB_1.
DR InterPro; IPR008138; SApB_2.
DR InterPro; IPR008140; SApB sub.
DR InterPro; IPR008373; SApOSin.
DR InterPro; IPR008139; SApOSinB.
DR Pfam; PF02199; SApA_2.
DR Pfam; PF05184; SApB_1; 3.
DR Pfam; PF03489; SApB_2; 4.
DR PRINTS; PR01797; SApOSin.
DR ProDom; PD001732; SApB sub; 3.
DR SMART; SM00162; SApA_2.
DR SMART; SM00118; SApB_4.
KM Signal; Glycoprotein; Repeat.
FT SIGNAL 1 16
FT CHAIN 17 554 SULFATED GLYCOPROTEIN 1.
FT DOMAIN 21 54 SApOSIN-LIKE TYPE A 1.
FT DOMAIN 61 138 SApOSIN-LIKE TYPE B 1.
FT DOMAIN 193 274 SApOSIN-LIKE TYPE B 2.
FT DOMAIN 310 391 SApOSIN-LIKE TYPE B 3.
FT DOMAIN 435 516 SApOSIN-LIKE TYPE B 4.
FT DOMAIN 521 554 SApOSIN-LIKE TYPE A 2.
FT DISULFID 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 197 270 BY SIMILARITY.
FT DISULFID 200 264 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 314 387 BY SIMILARITY.
FT DISULFID 317 381 BY SIMILARITY.
FT DISULFID 345 356 BY SIMILARITY.
FT DISULFID 439 512 BY SIMILARITY.
FT DISULFID 442 506 BY SIMILARITY.
FT DISULFID 470 481 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 115 115 P -> L (IN REF. 2 AND 3).
FT CONFLICT 299 299 D -> R (IN REF. 2).
FT CONFLICT 462 462 I -> V (IN REF. 3).
FT CONFLICT 527 527 W -> R (IN REF. 3).
FT CONFLICT 536 536 S -> M (IN REF. 3).
SQ SEQUENCE 554 AA; 61123 MW; DFE3F3AJA0520C6B CRC64;

Query Match 74.6%; Score 47; DB 1; Length 554;
Best Local Similarity 70.0%; Pred. No. 1.1;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 DVCQDCIQMV 11
Db 195 DVCQDCIKLV 204

RESULT 6
SAP_MOUSE STANDARD; PRT; 557 AA.
AC 061207; 060861; 064006; 064219;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN PSAP OR SGP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=92272718; PubMed=1590788;
RA Tsuda M., Sakiyama T., Endo H., Kitagawa T.;
RT "The primary structure of mouse saposin.";
RL Biochem. Biophys. Res. Commun. 184:1266-1272(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=94272317; PubMed=8003952;
RA Sprecher-Levy H., Orr-Urtreger A., Lonai P., Horowitz M.;
RT "Murine prosaposin: expression in the reproductive system of a gene
RL Cell. Mol. Biol. 40:233-233(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=96084310; PubMed=8565332;
RA Cao Q.P., Crain W.R.;
RT "Expression of SGP-1 mRNA in preimplantation mouse embryos.";
RL Dev. Genet. 17:263-271(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Zhao Q.Q., Hay N.N., Morales C.R.;
RT Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
RL -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: Contains 2 saposin A-type domains.
CC -1- SIMILARITY: Contains 4 saposin B-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announce/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; S36200; AAB22175.1; -
CC EMBL; S71616; AAB31059.1; -
CC EMBL; U27340; AAB32667.1; -
CC EMBL; U57999; AAB02695.1; -
CC PIR; JH0604; JH0604.

DR MGD; MGI:97783; Psap.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; Sapa_1.
DR InterPro; IPR008138; Sapa_2.
DR InterPro; IPR008140; Sapa_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008339; SaposinB.
DR Pfam; PF02199; Sapa_2.
DR Pfam; PF05184; Sapa_1; 4.
DR Pfam; PF03489; Sapa_2; 4.
DR PRINTS; PRO1797; SAPOSIN.
DR ProDom; PD001732; Sapa_sub; 3.
DR SMART; SM00152; SARA; 2.
DR SMART; SM00118; SARA; 4.
DR Signal; Glycoprotein; Repeat.
KM SIGNAL 1
FT CHAIN 17 557
FT DOMAIN 21 54
FT DOMAIN 59 142
FT DOMAIN 193 277
FT DOMAIN 313 394
FT DOMAIN 438 519
FT DOMAIN 524 557
FT DISULFID 63 138
FT DISULFID 66 132
FT DISULFID 94 106
FT DISULFID 197 273
FT DISULFID 200 267
FT DISULFID 229 240
FT DISULFID 317 390
FT DISULFID 320 384
FT DISULFID 348 359
FT DISULFID 442 515
FT DISULFID 445 509
FT DISULFID 473 484
FT CARBOHYD 80 80
FT CARBOHYD 214 214
FT CARBOHYD 334 334
FT CARBOHYD 459 459
FT CONFLICT 83 83
FT CONFLICT 158 158
FT CONFLICT 160 160
FT CONFLICT 171 172
FT CONFLICT 244 244
FT CONFLICT 254 254
FT CONFLICT 255 255
FT CONFLICT 260 262
FT CONFLICT 307 307
FT CONFLICT 322 322
FT CONFLICT 349 350
FT CONFLICT 367 367
FT CONFLICT 370 370
FT CONFLICT 373 373
FT CONFLICT 391 391
FT CONFLICT 393 393
FT CONFLICT 406 406
FT CONFLICT 430 430
FT CONFLICT 445 445
FT CONFLICT 448 448
SQ SEQUENCE 557 AA; 61422 MW; 134593E20499E35E CRC64;

Query Match 74.6%; Score 47; DB 1; Length 557;
Best Local Similarity 70.0%; Pred. No. 1.1;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 DVCQDCIQMV 11
Db 195 DVCQDCIKLV 204

RESULT 7
VENV_THOQV STANDARD; PRT; 512 AA.
ID VENV_THOQV
```


AC P28977;
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 GN Envelope glycoprotein precursor (Surface glycoprotein 75).
 DE P4.
 OS Thogoto virus (isolate Siar 126) (Tho).
 CC Virusess; asRNA negative-strand viruses; Orthomyxoviridae;
 CC Thogotovirinae.
 CC NCBI_TaxId=126796;
 CC (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92124738; PubMed=1733105;
 RA Morse M.A.; Marriott A.C.; Nuttall P.A.;
 RT "The glycoprotein of Thogoto virus (a tick-borne orthomyxo-like
 RL virus) is related to the baculovirus glycoprotein GP64.";
 CC Virology 186:640-646(1992).
 CC -1- FUNCTION: POSSIBLE ROLE IN ENDOCYTIC FUSION EVENTS DURING
 CC INFECTION.
 CC -1- SUBUNIT: Monomer (Probable).
 CC -1- SIMILARITY: TO DHORI VIRUS ENVELOPE GLYCOPROTEIN AND TO
 CC BACULOVIRUSES MAJOR ENVELOPE GLYCOPROTEIN (P64/P67).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, M77280; AAA47918.1; -;
 DR PIR; A40821; VGIYTH.
 DR InterPro; IPR004955; Baculo_gp64.
 DR Pfam; PF03273; Baculo_gp64; 1.
 DR KX Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 15
 FT CHAIN 16 512
 FT TRANSMEM 479 502
 FT CARBOHYD 185 185
 FT CARBOHYD 263 263
 FT CARBOHYD 289 289
 FT CARBOHYD 378 378
 FT CARBOHYD 416 416
 FT SEQUENCE 512 AA; 57550 MW; 0396FC36284A0DF1 CRC64;
 SQ
 Query Match 61.9%; Score 39; DB 1; Length 512;
 Best Local Similarity 71.4%; Pred. No. 23;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GPCVDCD 7
 DB 109 GDICDHC 115
 RESULT 8
 PCK5_BRACL STANDARD; PRT; 1696 AA.
 ID PCK5_BRACL Q9NJ15; Q9NJ16;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Proprietary convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)
 DE (Proprietary convertase PC6-like) (apc6).
 GN PC6.
 OS Branchiostoma californiensis (California lancelet) (Amphioxus).
 CC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 CC Branchiostoma
 CC NCBI_TaxId=7738;
 CC (1)
 RN SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
 RX MEDLINE=20175281; PubMed=10708868;
 RX Oliva A.A. Jr.; Chan S.J.; Steiner D.F.;

RT "Evolution of the prohormone convertases: identification of a
 RT homologue of PC6 in the protochordate amphioxus.";
 RT Biochim. Biophys. Acta 1477:338-348(2000).
 CC -1- FUNCTION: Likely to represent a widespread endoprotease activity
 CC within the constitutive and regulated secretory pathway. Capable
 CC of cleavage at the RX(K/R)R consensus motif (by similarity).
 CC -1- CATALYTIC ACTIVITY: Release of mature proteins from their
 CC propeptides by cleavage of Arg-Xaa-Yaa-Arg-|-Zaa bonds, where Xaa
 CC can be any amino acid and Yaa is Arg or Lys.
 CC -1- SUBCELLULAR LOCATION: ISOFORM A AND ISOFORM C ARE SECRETED.
 CC -1- ISOFORM B IS A TYPE I MEMBRANE PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=A;
 CC IsoId=Q9NJ15-1; Sequence=Displayed;
 CC Name=B;
 CC IsoId=Q9NJ15-2; Sequence=VSP_005444, VSP_005445;
 CC Name=C;
 CC IsoId=Q9NJ15-3; Sequence=VSP_005442, VSP_005443;
 CC -1- DOMAIN: The propeptide domain acts as an intramolecular chaperone
 CC assisting the folding of the zymogen within the endoplasmic
 CC reticulum.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC -1- SIMILARITY: Contains 1 homo B/P domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, AF184615; AAF26300.1; -;
 DR EMBL; AF184615; AAF26301.1; -;
 DR EMBL; AF184617; AAF26302.1; -;
 DR HSSP; Q99405; IMPT.
 DR MEROPS; S08 UPB; -;
 DR InterPro; IPR006212; Furin_repeat.
 DR InterPro; IPR009030; Grow_fac_recep.
 DR InterPro; IPR002029; Peptidase_S8.
 DR InterPro; IPR002884; Peptidase_S8b.
 DR InterPro; IPR009020; Protease_inhib.
 DR Pfam; PF01483; P_protein; 1.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR ProDom; PD000717; P_domain; 1.
 DR SMART; SMO0261; FU; 17.
 DR PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR HYDROLASE; Serine protease; Glycoprotein; Zymogen; Signal;
 KW Cleavage on pair of basic residues; Repeat; Alternative splicing;
 KW Transmembrane.
 FT SIGNAL 1 25
 FT PROPEP 26 110
 FT CHAIN 111 1696
 FT DOMAIN 111 1618
 FT TRANSMEM 1619 1639
 FT DOMAIN 1640 1696
 FT DOMAIN 111 488
 FT DOMAIN 496 637
 FT DOMAIN 664 1649
 FT SITE 110 111
 FT ACT_SITE 192 192
 FT ACT_SITE 233 233
 FT ACT_SITE 407 407
 FT CARBOHYD 246 246
 FT CARBOHYD 529 529
 FT CARBOHYD 885 885
 FT VARSPPLIC 1259 1323
 FT
 POTENTIAL.
 POTENTIAL.
 PROTEIN CONVERGENCE SUBTILISIN/KEXIN
 TYPE 5.
 EXTRACELLULAR (POTENTIAL).
 POTENTIAL.
 CYTOPLASMIC (POTENTIAL).
 CATALYTIC.
 HOMO B.
 CYS-RICH MOTIF (CRM) REGION.
 CLEAVAGE (AUTO-) (BY SIMILARITY).
 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CHARGE RELAY SYSTEM (BY SIMILARITY).
 CHARGE RELAY SYSTEM (BY SIMILARITY).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 DDTILDRGECTSCGPGYMBRREKCKACHPTECSDEY
 DDTACNDGFLTDASSCEAGCP -> AENONQNSFCPPA

```

FT FT PREVSVALAELALGHLARYSLDVPDQSNPPDYLGLADRRAL
FT FT TTATSAGRCA (in isoform C).
FT FT /FtId=VSP_005442.
FT FT Missing (in isoform C).
FT FT /FtId=VSP_005443.
FT FT CHPTCKCSDBYDDCTACNDGFLITDASCSCAGCPGQFL
FT FT HHGDCDCHRECKTC -> IARCDDRDRRCMDLVRFNC
FT FT VRRYFVRCCGCTCLYMEDRPMRGSSOPTQGRN (in
FT FT isoform A).
FT FT /FtId=VSP_005444.
FT FT Missing (in isoform A).
FT FT /FtId=VSP_005445.
SQ SEQUENCE 1696 AA; 188410 MW; 281CBE1784257CDB CRC64;

Query Match 61.9%; Score 39; DB 1; Length 1696;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DVCQDC 7
Db 775 DVCQDC 780

RESULT 9
CMG2_SCHPO STANDARD; PRT; 355 AA.
AC P32434;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Type I protein geranylgeranyltransferase beta subunit (EC 2.5.1.59)
DE (Type I protein geranylgeranyltransferase beta subunit) (GGTase-I-
DE beta) (PGGT).
GN CMG2 OR SPAC2E1P5.04C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetaceae;
NCBI_TaxID=4896;
[1]
RN NCB1
RP SEQUENCE FROM N.A.
RX MEDLINE=94085400; PubMed=8262067;
RA Diaz M., Sanchez Y., Bennett T., Dun C.R., Godoy C., Tamanoi F.,
RA Duran A., Perez P.;
RT "The Schizosaccharomycetes pombe cmg2+ gene codes for the beta subunit
RT of a geranylgeranyltransferase type I required for beta-glucan
RT synthesis.",
RL EMO J. 12:5245-5254(1993).
RN NCB1
RN NCB1
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voickart G., Aert R., Robben J., Glynnoprez B.,
RA Welljams I., Vansetrelle E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu S.J., Dreno S., Gloux S., Lelaure V., Mottier S.,
RA Lucas M., Rochet M., Galliard C., Tallade V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

```

```

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Uesary D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- FUNCTION: Catalyzes the transfer of a geranyl-geranyl moiety from
CC geranyl-geranyl pyrophosphate to proteins having the C-terminal C-
CC A-A-L where A is an aliphatic amino acid. In particular it
CC modifies the GTP-binding component of the 1,3-beta-D-glucan
CC synthase.
CC -1- CATALYTIC ACTIVITY: Geranyl[geranyl] diphosphate + protein-cysteine
CC = S-geranyl[geranyl]-protein + diphosphate.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit.
CC -1- SIMILARITY: Belongs to the protein prenyltransferase beta subunit
CC family.
CC -1- SIMILARITY: Contains 4 PFTB repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z12155; CA78143.1; -.
DR EMBL; AL163071; CAB86347.1; -.
DR PIR; S4186; S4186.
DR HSSP; Q02293; 1FT1.
DR GeneDB Spombe; SPAC2E1P5.04C; -.
DR InterPro; IPR001330; Prenyltrans.
DR InterPro; IPR008930; Terp_cyc_toroid.
DR Pfam; PF00432; Prenyltrans; 4.
KM Transferase; Prenyltransferase; Repeat; zinc.
FT REPEAT 119 163 PFTB 1.
FT REPEAT 169 210 PFTB 2.
FT REPEAT 234 275 PFTB 3.
FT REPEAT 282 324 PFTB 4.
FT METAL 260 260 PFTB 4.
FT METAL 262 262 ZINC (BY SIMILARITY).
FT METAL 312 312 ZINC (BY SIMILARITY).
SQ SEQUENCE 355 AA; 40023 MW; 811A3CBFF067071 CRC64;

Query Match 60.3%; Score 38; DB 1; Length 355;
Best Local Similarity 62.5%; Pred. No. 24;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 DVCQDCIQ 9
Db 343 NICRDCIQ 350

RESULT 10
Y777_TREPA STANDARD; PRT; 74 AA.
AC 083756;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0777.
GN TP0777.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OC NCB1_TaxID=160;
NCBI_TaxID=160;
[1]
RN NCB1
RP SEQUENCE FROM N.A.
RX STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RX Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RX Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RX Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khairat H., Richardson D., Howell J.K., Chidambaram M., Uitterback T.,

```

RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.,
RT "Complete genome sequence of *Treponema pallidum*, the syphilis
RT spirochete.";
RT Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AE001249; AAC65750.1; --
DR TIGR, B71282; B71282.
DR TIGR, TP0777; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 74 AA; 9018 MW; 299EB9E2CEC2D5 CRC64;

Query Match 58.7%; Score 37; DB 1; Length 74;
Best Local Similarity 50.0%; Pred. No. 7.7;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 DVCDCDQIMV 11
DB 29 DVCDCDQIMV 38

RESULT 11
NADA_HALNT1
ID NADA_HALNT1 STANDARD; PRT; 374 AA.
AC Q9H2N1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Quinolinate synthetase A.
GN NADA OR VNG1882G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasly S.R., Baliga N.S., Thorson V., Shroga J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welt R., Goo Y.A.,
RA Leitbauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Beck R.F., Pohlischer M., Spudis J.L., Jung K.-H.,
RA Alan M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RL
CC -!- FUNCTION: Catalyzes the condensation of iminoaspartate with
CC dihydroxyacetone phosphate to form quinolinate (By similarity).
CC -!- PATHWAY: NAD biosynthesis; aspartate to NADM; second step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the quinolinate synthetase A family.
CC Subfamily 3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AB005087; AAG20079.1; --
DR PIR, C84339; C84339.

DR HAMAP; MF 00569; -; 1.
DR InterPro; IPR003473; NADA.
DR Pfam; PF02445; NADA; 1.
DR TIGRFAMs; TIGR00550; nada; 1.
KW Pyridine nucleotide biosynthesis; Complete proteome.
SQ SEQUENCE 374 AA; 41305 MW; A52B4985169498B2 CRC64;

Query Match 58.7%; Score 37; DB 1; Length 374;
Best Local Similarity 60.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GDVCDQIMV 10
DB 321 GDVCDQIMV 330

RESULT 12
U183_HUMAN
ID U183_HUMAN STANDARD; PRT; 422 AA.
AC Q9BSU1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE UPF0183 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee Y.M., Kim W.;
RT "Identification of human homolog of rat lin-10.";
RT Submitted (Aug-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL
CC -!- SIMILARITY: Belongs to the UPF0183 family.
CC -!- CAUTION: Lin-10 has been incorrectly assigned for a number of
CC years, therefore what is stated in Ref.1 as a lin-10 homolog as
CC nothing to do with the real lin-10.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AF176088; AAK96888.1; --
DR EMBL, BC004556; AAH04556.1; --
DR InterPro, IPR00573; UPF0183.

DR Pfam; PF03676; UPF0183; 1.
SQ SEQUENCE 422 AA; 47524 MW; CEA42442E45A7E35 CRC64;

Query Match 58.7%; Score 37; DB 1; Length 422;
Best Local Similarity 54.5%; Pred. No. 41;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GDVQCDCIOMV 11
DB 242 GDSCQDVLMSL 252

RESULT 13

ID U183_MOUSE STANDARD; PRT; 422 AA.

AC 0922R1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE UPF0183 protein.
OC Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10990;
RN [1]

SEQUENCE FROM N.A.
MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Matovina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Schaeetz T.E.,
RA Brownstein M.U., Utsidi T.B., Toshiyuki S., Carninci P., Prange S.J.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultix S.W.,
RA Villalón D.K., Murzy D.M., Sodergren E.U., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SIMILARITY: Belongs to the UPF0183 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; BC006874; AAH06874.1;
DR MGD; MGI:2384564; BC006874;
DR InterPro; IPR005373; UPF0183.
DR Pfam; PF03676; UPF0183; 1.
SQ SEQUENCE 422 AA; 47402 MW; 5639E30A26D96F14 CRC64;

Query Match 58.7%; Score 37; DB 1; Length 422;
Best Local Similarity 54.5%; Pred. No. 41;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GDVQCDCIOMV 11
DB 242 GDSCQDVLMSL 252

RESULT 14

U183_RAT STANDARD; PRT; 422 AA.

AC 008654;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE UPF0183 protein.

OC Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley;
RA MEDLINE=96153816; PubMed=9480860;

RA Ide N., Hirao K., Hata Y., Takeuchi M., Irie M., Yao I., Deguchi M.,
RA Toyoda A., Nishiohara H., Mizoguchi A., Takai Y.,
RT "Molecular cloning and characterization of rat ltn-10.";
RL Biochem. Biophys. Res. Commun. 243:634-638(1998).
CC -1- SIMILARITY: Belongs to the UPF0183 family.
CC -1- CAUTION: ltn-10 has been incorrectly assigned for a number of
CC years, therefore what is stated in Ref.1 as a ltn-10 homolog as
CC nothing to do with the real ltn-10.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U92010; AAB51383.1;
DR PIR; JE0239; J02039.
DR InterPro; IPR005373; UPF0183.
DR Pfam; PF03676; UPF0183; 1.
SQ SEQUENCE 422 AA; 47463 MW; D5E185189FE2E320 CRC64;

Query Match 58.7%; Score 37; DB 1; Length 422;
Best Local Similarity 54.5%; Pred. No. 41;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GDVQCDCIOMV 11
DB 242 GDSCQDVLMSL 252

RESULT 15

CLPE_LACLA STANDARD; PRT; 748 AA.

AC 09C109;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP-dependent Clp protease ATP-binding subunit clpE.

CLPE OR L10557.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RA MEDLINE=21235186; PubMed=11337471;

RA Bolotin A., Wincker P., Manger S., Jallion O., Malarme K.,
RA Weisenbach J., Ehrlich S.D., Sorokin A.,
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis sep. lactis il1403.";
RL Genome Res. 11:731-753(2001).
CC -1- FUNCTION: COULD BE NECESSARY FOR DEGRADING PROTEINS GENERATED BY
CC CERTAIN TYPES OF STRESS (By similarity).
CC -1- INDUCTION: By heat shock (Probable).
CC -1- SIMILARITY: BELONGS TO THE CLPA/CLPB FAMILY. CLPE SUBFAMILY.

CC -1- SIMILARITY: Contains 1 UVR domain.

```
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006289; AK04655.1; -.
DR PIR; E86694; E86694.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003599; AAA ATPase centr.
DR InterPro; IPR001270; Chaperin_c1pa/B.
DR InterPro; IPR001943; UvrB/C.
DR Pfam; PF00004; AAA; 2.
DR Pfam; PF02151; UVR; 1.
DR PRINTS; PR00300; CLPPROTEASEA.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00870; CLPAB_1; 1.
DR PROSITE; PS00871; CLPAB_2; 1.
DR PROSITE; PS50151; UVR; 1.
DR Chaperone; ATP-binding; Heat shock; Zinc-finger; Repeat;
KW Complete proteome.
KW DOMAIN 107 354 I.
FT DOMAIN 359 394 UVR.
FT DOMAIN 407 598 II.
FT ZN_FING 3 32 C4-TYPE (POTENTIAL).
FT NP_BIND 152 159 ATP (POTENTIAL).
FT NP_BIND 481 488 ATP (POTENTIAL).
SQ SEQUENCE 748 AA; 83144 MW; 738ECF1B994B31B7 CRC64;

Query Match 58.7%; Score 37; DB 1; Length 748;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 DVCQDCIQMV 11
DB 27 DLQNCYQIM 36
```

Search completed: May 5, 2004, 13:31:44
Job time : 1.27161 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:26:23 ; Search time 0.438746 Seconds
(without alignments)
2411.659 Million cell updates/sec

Title: US-09-743-684A-19
Perfect score: 63
Sequence: 1 GDVQCDCIQMW 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	100.0	79	2 A49475	cerebroside sulfat
2	63	100.0	527	1 SAHUP	saposin precursor
3	47	74.6	554	1 A28716	saposin precursor
4	47	74.6	557	1 JH0604	saposin precursor
5	40	63.5	46	2 E69830	hypothetical prote
6	40	63.5	1356	2 F84486	probable retroel
7	39	61.9	320	2 B49810	gene 7 protein - m
8	39	61.9	512	1 VGIVTH	envelope glycoprot
9	39	61.9	666	2 T10567	probable serine/th
10	38	60.3	221	2 T46918	hypothetical prote
11	38	60.3	307	2 T40043	probable transport
12	38	60.3	320	2 E70355	conserved hypothet
13	38	60.3	348	2 A57234	lin-44 protein pre
14	38	60.3	355	2 S41686	geranylgeranyltran
15	38	60.3	456	2 T24442	hypothetical prote
16	38	60.3	1672	2 T46237	hypothetical prote
17	38	60.3	2609	2 T40399	probable transport
18	37	58.7	74	2 B71282	hypothetical prote
19	37	58.7	374	2 C84339	quinolinolate synthe
20	37	58.7	413	2 F86847	hypothetical prote
21	37	58.7	422	2 J80239	lin-10 protein - r
22	37	58.7	429	2 T41560	phenylalanyl-trna
23	37	58.7	748	2 E86694	ATP-dependent prot
24	37	58.7	5147	1 IJFFTM	cadherin-related t
25	36	57.1	47	2 I48943	cellular disintegr
26	36	57.1	111	2 B71239	hypothetical prote
27	36	57.1	151	2 S60259	meltrin gamma - mo
28	36	57.1	213	2 F70476	conserved hypothet
29	36	57.1	218	2 AH1388	late competence pr

30	36	57.1	291	2 T20666	hypothetical prote
31	36	57.1	294	2 T20664	hypothetical prote
32	36	57.1	330	2 A29915	teratocarcinoma gl
33	36	57.1	367	2 T24298	hypothetical prote
34	36	57.1	493	2 T29833	hypothetical prote
35	36	57.1	511	2 S10527	endoglucanase B pr
36	36	57.1	655	2 JC7850	disintegrin and me
37	36	57.1	660	2 S71949	metalloproteinase
38	36	57.1	778	2 D85055	probable polypro
39	36	57.1	1363	2 T47492	copia-like polypro
40	35.5	56.3	1416	2 E88550	protein ZC84.1 [im
41	35.5	56.3	2844	2 S28291	hypothetical prote
42	35	55.6	101	2 S62337	L7i-8 protein - fr
43	35	55.6	103	2 A47337	rhodniin - Rhodniu
44	35	55.6	108	2 A86675	hypothetical prote
45	35	55.6	127	2 F87303	cheU protein [impo

ALIGNMENTS

RESULT 1

A49475

cerebroside sulfat

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 17-May-1996

C:Accession: A49475

R:Stevens, R.L.; Faull, K.F.; Conklin, K.A.; Green, B.N.; Fluharty, A.L.

Biochemistry 32, 4051-4059, 1993

A:Title: Porcine cerebroside sulfate activator: further structural characterization and

A:Reference number: A49475; MUID:93229506; PMID:8471613

A:Accession: A49475

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-79 <STR>

A:Experimental source: kidney

A:Note: sequence extracted from NCBI backbone (NCBIP:129597)

C:Superfamily: saposin; saposin repeat homology

F:1-79/Domain: saposin repeat homology <SAP>

Query Match 100.0%; Score 63; DB 2; Length 79;

Best Local Similarity 100.0%; Pred. No. 0.0019;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMW 11

Db 1 GDVQCDCIQMW 11

RESULT 2

SAHUP

saposin precursor [validated] - human

N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component

ein (SAP); sphingolipid activator protein A2; sulfate sulfatase activator protein

N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D

C:Species: Homo sapiens (man)

C:Date: 30-Jun-1992 #sequence_revision 17-Nov-1995 #text_change 08-Dec-2000

C:Accession: JX0061; A57368; A42003; C42003; D42003; A30367; S34740; S36140; S36

0226; I37265; I37264

R:Nakano, T.; Sandhoff, K.; Stuenkel, J.; Christomanou, H.; Suzuki, K.

J. Biochem. 105, 152-154, 1989

A:Title: Structure of full-length cDNA coding for sulfate activator, a Co-beta-glucosi

A:Reference number: JX0061; MUID:89255151; PMID:2498298

A:Accession: JX0061

A:Molecule type: mRNA

A:Residues: 1-527 <NAK>

A:Cross-references: GB:P00422; NID:g220063; PIDN:BAA00321.1; PID:g220064

A:Note: alternative splice form 1

A:Accession: A57368

A:Molecule type: mRNA

A:Residues: 1-259,263-527 <NA2>

A:Cross-references: GB:J03015; GB:J03086; NID:g337755; PIDN:AAB59494.1; PID:g337756

A:Note: alternative splice form 2

R;Rorman, E.G.; Scheinker, V.; Grabowski, G.A.
 Genomics 13, 312-318, 1992
 A;Title: Structure and evolution of the human prosaposin chromosomal gene.
 A;Reference number: A42003; MUID:92307663; PMID:1612590
 A;Accession: A42003
 A;Molecule type: DNA
 A;Residues: 50-140 <ROR>
 A;Cross-references: GB:M86181
 A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107236)
 A;Accession: B42003
 A;Molecule type: DNA
 A;Residues: 185-259;263-276 <RO2>
 A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107237)
 A;Accession: C42003
 A;Molecule type: DNA
 A;Residues: 305-393 <RO3>
 A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107238); sequence inc
 A;Accession: D42003
 A;Molecule type: DNA
 A;Residues: 399-487 <RO4>
 A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107239); sequence inc
 R;Rorman, E.G.; Grabowski, G.A.
 Genomics 5, 486-492, 1989
 A;Title: Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphin
 A;Reference number: A30367; MUID:90129043; PMID:2515150
 A;Accession: A30367
 A;Molecule type: mRNA
 A;Residues: 1-259;263-527 <RO5>
 A;Cross-references: GB:J03077; NID:G1832230; PIDN:AAA52560.1; PID:G1832231
 A;Note: alternative splice form 2
 R;Hiraiwa, M.; O'Brien, J.S.; Kishimoto, Y.; Gaidzicka, M.; Fluharty, A.L.; Gimms, E.I.;
 Arch. Biochem. Biophys. 304, 110-116, 1993
 A;Title: Isolation, characterization, and proteolysis of human prosaposin, the precursor
 A;Reference number: S34740; MUID:93311991; PMID:8323276
 A;Accession: S34740
 A;Molecule type: protein
 A;Residues: 17-24;165-172;180-189;301-305 <HTR>
 R;Tynnela, J.; Palmer, D.N.; Baumann, M.; Haltia, M.
 FEBS Lett. 330, 8-12, 1993
 A;Title: Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis.
 A;Reference number: S36140; MUID:93380576; PMID:8370464
 A;Accession: S36140
 A;Molecule type: protein
 A;Residues: 'XX', 62, 'X', 64-65, 'X', 67-79, 'X', 81-84 <TYV>
 A;Note: saposin A
 A;Accession: S36141
 A;Molecule type: protein
 A;Residues: 'XXX', 413-414, 'X', 416-428, 'X', 430-434 <TY2>
 R;Holtzman, H.; Sandhoff, K.; Kwon, H.Y.; Harzer, K.; Nakano, T.; Suzuki, K.
 J. Biol. Chem. 266, 7556-7560, 1991
 A;Title: Sulfatide activator protein. Alternative splicing that generates three mRNAs an
 A;Reference number: S36988; MUID:91210267; PMID:2019586
 A;Accession: S36988
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-240, 'S', 242-527 <HO1>
 A;Cross-references: EMBL:M60255; NID:G337759; PIDN:AAA36594.1; PID:G337760
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
 A;Title: cerebroside sulfate activator protein mutant MU-9; corresponds to alternative sp
 A;Accession: S36989
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-240, 'S', 242-259, 263-527 <HO2>
 A;Cross-references: EMBL:M60257; NID:G337764; PIDN:AAA36595.1; PID:G337765
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
 A;Title: cerebroside sulfate activator protein mutant MU-0; corresponds to alternative sp
 A;Accession: S36990
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-240, 'S', 242-259, 261-527 <HO3>
 A;Cross-references: EMBL:M60258; NID:G337766; PIDN:AAA36596.1; PID:G337767
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991

A;Note: cerebroside sulfate activator protein mutant MU-6; corresponds to alternative sp
 R;Kondo, K.; Hinano, T.; Sano, A.; Kakimoto, Y.
 Biochem. Biophys. Res. Commun. 181, 286-292, 1991
 A;Title: Isolation and characterization of prosaposin from human milk.
 A;Reference number: PS0330; MUID:92068206; PMID:1958198
 A;Accession: PS0330
 A;Molecule type: protein
 A;Residues: 17-24, 'X', 26 <KON>
 A;Experimental source: milk
 R;Kretz, K.A.; Carson, G.S.; Morimoto, S.; Kishimoto, Y.; Fluharty, A.L.; O'Brien, J.S.
 Proc. Natl. Acad. Sci. U.S.A. 87, 2541-2544, 1990
 A;Title: Characterization of a mutation in a family with saposin B deficiency: a glycosyl
 A;Reference number: A35985; MUID:90207231; PMID:2320574
 A;Accession: A35985
 A;Molecule type: mRNA
 A;Residues: 213-221 <KRE>
 A;Cross-references: GB:M32221
 A;Accession: B35985
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-259;263-527 <KR2>
 A;Cross-references: GB:M32221; NID:G337761; PIDN:AAA60303.1; PID:G337762
 A;Experimental source: lymphoblast
 A;Accession: C35985
 A;Molecule type: mRNA
 A;Residues: 213-216, 'I', 218-221 <KR3>
 A;Note: sequence from patients with activator-deficient metachromatic leukodystrophy; th
 R;Fuerst, W.; Schubert, J.; Machleidt, W.; Meyer, H.E.; Sandhoff, K.
 Eur. J. Biochem. 192, 709-714, 1990
 A;Title: The complete amino-acid sequences of human ganglioside GM2 activator protein an
 A;Reference number: S13195; MUID:91006165; PMID:2209618
 A;Accession: S13196
 A;Molecule type: protein
 A;Residues: 195-259, 263-277 <FUE>
 R;Morimoto, S.; Martin, B.M.; Yamamoto, Y.; Kretz, K.A.; O'Brien, J.S.; Kishimoto, Y.
 Proc. Natl. Acad. Sci. U.S.A. 86, 3389-3393, 1989
 A;Title: Saposin A: second cerebroside activator protein.
 A;Reference number: A32784; MUID:89240739; PMID:2717620
 A;Accession: A32784
 A;Molecule type: protein
 A;Residues: 60-84;86-107;109-119;125-134 <MOR>
 R;O'Brien, J.S.; Kretz, K.A.; Dewji, N.; Wenger, D.A.; Esch, F.; Fluharty, A.L.
 Science 241, 1098-1101, 1988
 A;Title: Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic
 A;Reference number: A41240; MUID:88321660; PMID:2842863
 A;Accession: A41240
 A;Molecule type: mRNA
 A;Residues: 'GSSR', 18-259, 263-299, 'D', 301-302, 'D', 304-527 <ORAB>
 A;Cross-references: GB:J03086
 R;Dewji, N.N.; Wenger, D.A.; O'Brien, J.S.
 Proc. Natl. Acad. Sci. U.S.A. 84, 8652-8656, 1987
 A;Title: Nucleotide sequence of cloned cDNA for human sphingolipid activator protein 1 p
 A;Reference number: S02289; MUID:88068647; PMID:2825202
 A;Accession: S02289
 A;Status: significant sequence differences
 A;Molecule type: mRNA
 A;Cross-references: EMBL:J03015
 A;Note: this sequence corrected by A41240
 A;Note: part of this sequence, including the amino end of the mature protein, was determ
 R;Klein Schmidt, T.; Christomanou, H.; Braunitzer, G.
 Biol. Chem. Hoppe-Seyler 369, 1361-1365, 1988
 A;Title: Complete amino-acid sequence of the naturally occurring A(2) activator protein f
 A;Reference number: S02028; MUID:89207118; PMID:3242555
 A;Accession: S02028
 A;Molecule type: protein
 A;Residues: 195-259, 263-276 <KLE>
 R;Fuerst, W.; Machleidt, W.; Sandhoff, K.
 Biol. Chem. Hoppe-Seyler 369, 317-328, 1988
 A;Title: The precursor of sulfatide activator protein is processed to three different pr
 A;Reference number: S00813; MUID:89000190; PMID:3048308
 A;Accession: S00813
 A;Molecule type: protein
 A;Residues: 410-487 <FU2>

R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G.
Biol. Chem. Hoppe-Seyler 368, 1571-1578, 1987
A;Title: Complete amino-acid sequence and carbohydrate content of the naturally occurring
A;Reference number: S00226; MUID:88163077; PMID:3442600
A;Accession: S00226
A;Molecule type: protein
A;Residues: 314-393 <KL2>
R;Vaccaro, A.M.; Salvio, R.; Barca, A.; Tatti, M.; Ciaffoni, F.; Maras, B.; Siciliano,
J. Biol. Chem. 270, 9953-9960, 1995
A;Title: Structural analysis of saposin C and B. Complete localization of disulfide bridge
A;Reference number: A57297; MUID:95247790; PMID:7730378
A;Contents: annotation; disulfide bonds; glycosylation
R;Holtzman, H.; Sandhoff, K.; Fuerst, W.; Kwon, H.Y.; Schnabel, D.; Suzuki, K.
FEBS Lett. 280, 267-270, 1991
A;Title: The organization of the gene for the human cerebroside sulfate activator protein
A;Reference number: I37264; MUID:91192146; PMID:2013321
A;Accession: I37265
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 59-125 <RES>
A;Cross-references: EMBL:X57107; NID:G30234; PIDN:CAA40391.1; PID:G30235
A;Accession: I37264
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 307-516 <RE2>
A;Cross-references: EMBL:X57108; NID:G30232; PIDN:CAA40392.1; PID:G1565257
A;Note: sequence revised relative to PID:G30233 (corrected coding region)
C;Genetics:
A;Gene: GDB:PSAP; GLRA
A;Cross-references: GDB:120366; OMIM:176801
A;Map position: 10q22.1-10q22.1
A;Introns: 83/3; 338/3; 401/1; 453/3; 480/3
A;Note: defects in this gene may cause variant Gaucher disease, variant Tay-Sachs disease
A;Note: list of introns is incomplete

Query Match 100.0%; Score 63; DB 1; Length 527;
Best Local Similarity 100.0%; Pred. No. 0.0091;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11
|||||
Db 195 GDVQCDCIQMV 205
|||||

RESULT 3
A28716
saposin precursor - rat
N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfate
N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A28716
R;Collard, M.W.; Sylvester, S.R.; Tsuruta, J.K.; Griswold, M.D.
Biochemistry 27, 4557-4564, 1988
A;Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat
A;Reference number: A28716; MUID:89000647; PMID:3048385
A;Accession: A28716
A;Molecule type: mRNA
A;Residues: 1-554 <COL>
A;Cross-references: GB:M19936; NID:G206904; PIDN:AAA42136.1; PID:G206905
A;Note: parts of this sequence, including the amino end of the mature protein, were deter
C;Function:
A;Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc
A;Pathway: sphingolipid catabolism
A;Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy
A;Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsul
A;Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiester
C;Superfamily: saposin; saposin repeat homology
C;Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-554/Product: prosaposin #status predicted <PRO>

F;60-143/Product: saposin A #status predicted <SAPA>
F;189-280/Domain: saposin B #status predicted <SAP2>
F;194-273/Product: saposin B #status predicted <SAB1>
F;306-397/Domain: saposin repeat homology <SAP3>
F;310-389/Product: saposin C #status predicted <SAPC>
F;431-522/Domain: saposin repeat homology <SAP4>
F;437-514/Product: saposin D #status predicted <SAPD>
F;63-138,66-132,94-106,439-512,442-506,470-481/Disulfide bonds: #status predicted
F;80,214,331,456/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;197-270,200-264,229-240,314-387,317-381,345-356/Disulfide bonds: #status predicted

Query Match 74.6%; Score 47; DB 1; Length 554;
Best Local Similarity 70.0%; Pred. No. 3.6;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DVCQDCIQMV 11
|||||
Db 195 DVCQDCIKLV 204
|||||

RESULT 4
JH0604
saposin precursor - mouse
N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfate
N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: JH0604
R;Tsuda, M.; Sakiyama, T.; Endo, H.; Kitagawa, T.
Biochem. Biophys. Res. Commun. 184, 1266-1272, 1992
A;Title: The primary structure of mouse saposin.
A;Reference number: JH0604; MUID:92272718; PMID:1590788
A;Accession: JH0604
A;Molecule type: mRNA
A;Residues: 1-557 <TSU>
A;Cross-references: GB:S36200; NID:G249386; PIDN:AA22175.1; PID:G249387
A;Experimental source: liver
C;Function:
A;Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc
A;Pathway: sphingolipid catabolism
A;Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy
A;Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsul
A;Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiester
C;Superfamily: saposin; saposin repeat homology
C;Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-557/Product: prosaposin #status predicted <PRO>
F;55-148/Domain: saposin repeat homology <SAP1>
F;60-143/Product: saposin A #status predicted <SAPA>
F;189-283/Domain: saposin repeat homology <SAP2>
F;194-276/Product: saposin B #status predicted <SAB1>
F;309-400/Domain: saposin repeat homology <SAP3>
F;313-392/Product: saposin C #status predicted <SAPC>
F;434-525/Domain: saposin repeat homology <SAP4>
F;440-517/Product: saposin D #status predicted <SAPD>
F;63-138,66-132,94-106,197-273,200-267,229-240,317-390,320-384,348-359,442-515,445-509,4
F;80,214,334,379,459/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 74.6%; Score 47; DB 1; Length 557;
Best Local Similarity 70.0%; Pred. No. 3.6;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DVCQDCIQMV 11
|||||
Db 195 DVCQDCIKLV 204
|||||

RESULT 5
E69830
hypothetical protein yfH - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999

C:Accession: E69830
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertea
 C.; Bron, S.; Broutillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
 Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: E69830
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-46 <KUN>
 A:Cross-references: GB:Z99109; GB:AL009126; NID:G2633260; PIDN:CAB12863.1; PID:e1183025;
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yhfH

Query Match 63.5%; Score 40; DB 2; Length 46;
 Best Local Similarity 50.0%; Pred. No. 6.1;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 GDVQDCIQM 10
 |||||
 Db 34 GNICNDCKV 43

RESULT 6
 F84486
 Probable retroelement pol polyprotein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: F84486
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: F84486
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1356 <STO>
 A:Cross-references: GB:AE002093; NID:G4895171; PIDN:AAD32759.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: Atg07550
 A:Map position: 2
 C:Superfamily: retrovirus-related polyprotein

Query Match 63.5%; Score 40; DB 2; Length 1356;
 Best Local Similarity 85.7%; Pred. No. 1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DVQDCI 8
 |||||
 Db 471 DVCEDCI 477

RESULT 7
 B49810
 gene 7 protein - murine rotavirus B
 C:Species: group B rotavirus
 A:Note: strain IDIR (infectious diarrhea of infant rats)
 C:Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 18-Jul-2001
 C:Accession: B49810

R:Eiden, J.J.
 Virology 199, 212-218, 1994
 A:Title: Expression and sequence analysis of gene 7 of the IDIR agent (group B rotavirus)
 A:Reference number: A49810; MUID:94160571; PMID:8116245
 A:Accession: B49810
 A:Status: preliminary
 A:Molecule type: genomic RNA
 A:Residues: 1-320 <EID>
 A:Cross-references: GB:U01164; NID:G470378; PIDN:AAA60453.1; PID:G470379
 C:Superfamily: rotavirus B gene 7 protein

Query Match 61.9%; Score 39; DB 2; Length 320;
 Best Local Similarity 50.0%; Pred. No. 45;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 GDVQDCIQM 10
 |||||
 Db 157 GDKQCNCLRL 166

RESULT 8
 VGIIVTH
 envelope glycoprotein - Thogoto virus (strain SiAr126)
 N:Alternate names: surface glycoprotein 75
 C:Species: Thogoto virus
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
 C:Accession: A40821
 R:Morse, M.A.; Marriott, A.C.; Nuttall, P.A.
 Virology 186, 640-646, 1992
 A:Title: The glycoprotein of Thogoto virus (a tick-borne orthomyxo-like virus) is related
 A:Reference number: A40821; MUID:92124738; PMID:1733105
 A:Accession: A40821
 A:Molecule type: genomic RNA
 A:Residues: 1-512 <WOR>
 A:Cross-references: GB:M77280; NID:G335213; PIDN:AAA47918.1; PID:G335214
 C:Genetics:
 A:Map position: segment 4
 C:Superfamily: baculovirus major envelope glycoprotein
 C:Keywords: envelope protein; glycoprotein; transmembrane protein
 F:185,263,289,378,416/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 61.9%; Score 39; DB 1; Length 512;
 Best Local Similarity 71.4%; Pred. No. 67;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GDVQDC 7
 |||||
 Db 109 GDICHDC 115

RESULT 9
 TI0567
 Probable serine/threonine-specific protein kinase (EC 2.7.1.-) F2584.90 - Arabidopsis th
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jan-2000
 C:Accession: TI0567
 R:Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft
 submitted to the Protein Sequence Database, June 1999
 A:Reference number: Z16533
 A:Accession: TI0567
 A:Molecule type: DNA
 A:Residues: 1-666 <BEV>
 A:Cross-references: EMBL:AL050399; GSPDB:GN00062; ATSP:F25E4.90
 A:Experimental source: cultivar Columbia; BAC clone F25E4
 C:Genetics:
 A:Gene: ATSP:F25E4.90
 A:Map position: 4
 A:Introns: 315/1; 355/3; 426/1; 505/2; 556/3
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase
 C:Keywords: phosphotransferase; protein kinase
 F:337-618/Domain: protein kinase homology <KIN>

Query Match 61.9%; Score 39; DB 2; Length 666;

Best Local Similarity 75.0%; Pred. No. 83;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VCQDCIQM 10
Db 86 VCSDCIQL 93

RESULT 10
T46918
hypothetical protein DKFZp762L137.1 - human
C:Species: Homo sapiens (man)
C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 20-Apr-2000
C:Accession: T46918
R:Ottenwälder, B.; Obermaier, B.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, February 2000
A:Reference number: Z24136
A:Accession: T46918
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-221 <AAA>
A:Cross-references: EMBL:AL157432
A:Experimental source: adult melanoma (MeWo cell line); clone DKFZp762L137
C:Genetics:
A:Note: DKFZp762L137.1
C:Superfamily: human hypothetical protein DKFZp762L137.1

Query Match 60.3%; Score 38; DB 2; Length 221;
Best Local Similarity 45.5%; Pred. No. 48;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11
Db 50 GDICNACVLLV 60

RESULT 11
T40043
probable transport protein - fission yeast (Schizosaccharomyces pombe) (fragment)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T40043
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z21901
A:Accession: T40043
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-307 <WOO>
A:Cross-references: EMBL:AL031517; PIDN:CAA20653.1; GSPDB:GN00067; SPDB:SPBC28E12.06c
A:Experimental source: strain 972h-; cosmid c28E12
C:Genetics:
A:Gene: spbc3h7.16; SPDB:SPBC28E12.06c
A:Map position: 2

Query Match 60.3%; Score 38; DB 2; Length 307;
Best Local Similarity 50.0%; Pred. No. 63;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 DVCQDCIQMV 11
Db 277 NVCSDCVSM 286

RESULT 12
E70355
conserved hypothetical protein aq_622 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: E70355
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: E70355
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-320 <AQF>
A:Cross-references: GB:AE000698; NID:g2983224; PIDN:AAC06834.1; PID:g2983237; GB:AE00065
A:Experimental source: strain VF5
C:Genetics:
A:Gene: aq_622

Query Match 60.3%; Score 38; DB 2; Length 320;
Best Local Similarity 60.0%; Pred. No. 65;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 DVCQDCIQMV 11
Db 179 EVCLDCIDML 188

RESULT 13
A57234
lin-44 protein precursor - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 20-Sep-1999
C:Accession: A57234
R:Herman, M.A.; Vassilieva, L.L.; Horvitz, H.R.; Shaw, J.E.; Herman, R.K.
Cell 83, 101-110, 1995
A:Title: The Caenorhabditis elegans gene lin-44, which controls the polarity of certain
submitted to the EMBL Data Library, September 1995
A:Reference number: A57234; MUID:96006529; PMID:7553861
A:Accession: A57234
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-348 <HER>
A:Cross-references: GB:U22179; NID:g758776; PID:g758777
C:Superfamily: int-1 transforming protein

Query Match 60.3%; Score 38; DB 2; Length 348;
Best Local Similarity 85.7%; Pred. No. 70;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VCQDCIQ 9
Db 325 VCKDCIQ 331

RESULT 14
S41686
geranylgeranyltransferase type I (EC 2.5.1.-) beta chain - fission yeast (Schizosaccharo
C:Species: Schizosaccharomyces pombe
C:Date: 31-Mar-1992 #sequence_revision 14-Sep-1994 #text_change 21-Jan-2000
C:Accession: S41686
R:Diaz, M.; Sanchez, Y.; Bennett, T.; Sun, C.R.; Godoy, C.; Tamanoi, F.; Duran, A.; Pere
EMBO J. 12, 5245-5254, 1993
A:Title: The Schizosaccharomyces pombe cwg2(+) gene codes for the beta subunit of a gera
A:Reference number: S41686; MUID:94085400; PMID:8262067
A:Accession: S41686
A:Molecule type: DNA
A:Residues: 1-355 <DIA>
A:Cross-references: EMBL:Z12155; NID:g4946; PIDN:CAA78143.1; PID:g396477
C:Genetics:
A:Gene: cwg2
C:Superfamily: cell division control protein CDC43
C:Keywords: glycoprotein; transferase; transmembrane protein
F:31-47/Domain: transmembrane #status predicted <TM1>
F:95-111/Domain: transmembrane #status predicted <TM2>
F:113,215,219,222/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 60.3%; Score 38; DB 2; Length 355;
Best Local Similarity 62.5%; Pred. No. 71;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DVCQDCIQ 9

```

Db          343 NICKDCIQ 350
:|:|:|
RESULT 15
T24442
hypothetical protein T04B2.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Apr-2003
C;Accession: T24442; T26042
R;Coles, L.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z19890
A;Accession: T24442
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-456 <WIL>
A;Cross-references: EMBL:Z68299; PIDN:CAA92615.1; GSPDB:GN00022; CESP:T04B2.5
A;Experimental source: clone T04B2
R;Coles, L.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z20141
A;Accession: T26042
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-456 <W12>
A;Cross-references: EMBL:Z68301; PIDN:CAA92628.1; GSPDB:GN00022; CESP:T04B2.5
A;Experimental source: clone W01B6
C;Genetics:
A;Gene: CESP:T04B2.5
A;Map position: 4
A;Introns: 48/3; 104/3; 139/1; 424/3
C;Superfamily: patatin

Query Match      60.3%; Score 38; DB 2; Length 456;
Best Local Similarity 50.0%; Pred. No. 88;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 GDVQCDCIQ 10
      | | | | | | | |
Db      97 GSVCECAEM 106

Search completed: May 5, 2004, 13:34:35
Job time : 0.438746 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:14:53 ; Search time 1.16999 Seconds
(without alignments)
2606.077 Million cell updates/sec

Title: US-09-743-684A-19
Perfect score: 63
Sequence: 1 GDVQCDCIQMV 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	100.0	209	14	US-10-043-487-340
2	63	100.0	479	10	US-09-978-418-40
3	63	100.0	523	9	US-09-767-007A-2
4	63	100.0	524	9	US-09-870-759-60
5	63	100.0	524	10	US-09-751-708A-60
6	63	100.0	524	12	US-10-267-502-386
7	63	100.0	527	9	US-09-870-759-61
8	63	100.0	527	10	US-09-751-708A-61
9	63	100.0	527	14	US-10-060-036-73
10	57	90.5	156	12	US-09-925-298-644
11	57	90.5	156	14	US-10-102-806-644
12	47	74.6	554	14	US-10-205-194-176
13	42	66.7	521	15	US-10-276-162-1
14	42	66.7	531	11	US-09-833-245-903
15	41	65.1	60	14	US-10-083-357-929

16	40	63.5	22	14	US-10-300-694A-73	Sequence 73, Appl
17	40	63.5	532	14	US-10-128-714-3442	Sequence 3442, Ap
18	40	63.5	532	14	US-10-128-714-8442	Sequence 8442, Ap
19	39	61.9	48	12	US-10-424-599-148377	Sequence 148377,
20	39	61.9	171	12	US-10-424-599-269800	Sequence 269800,
21	39	61.9	362	12	US-10-332-426-8	Sequence 8, Appli
22	39	61.9	653	14	US-10-081-872-194	Sequence 194, App
23	39	61.9	653	15	US-10-385-305-194	Sequence 194, App
24	38	60.3	166	12	US-10-425-114-68084	Sequence 68084, A
25	38	60.3	221	12	US-10-170-385-441	Sequence 441, App
26	38	60.3	251	14	US-10-106-698-4737	Sequence 4737, Ap
27	38	60.3	252	9	US-09-925-301-993	Sequence 993, App
28	38	60.3	469	12	US-10-282-122A-50094	Sequence 50094, A
29	37	58.7	110	9	US-09-764-864-848	Sequence 848, App
30	37	58.7	129	12	US-10-424-599-249703	Sequence 249703,
31	37	58.7	166	12	US-10-424-599-277844	Sequence 277844,
32	37	58.7	178	9	US-09-764-864-1303	Sequence 1303, Ap
33	37	58.7	319	9	US-09-764-868-659	Sequence 659, App
34	37	58.7	333	12	US-10-243-552-875	Sequence 875, App
35	37	58.7	365	13	US-10-029-180-62	Sequence 62, Appl
36	37	58.7	429	15	US-10-369-493-22639	Sequence 22639, A
37	37	58.7	465	14	US-10-103-313-357	Sequence 357, App
38	37	58.7	530	12	US-10-282-122A-74259	Sequence 74259, A
39	37	58.7	595	15	US-10-369-493-11077	Sequence 11077, A
40	37	58.7	748	15	US-10-369-493-18341	Sequence 18341, A
41	37	58.7	760	12	US-10-282-122A-74644	Sequence 74644, A
42	37	58.7	766	15	US-10-094-749-3040	Sequence 3040, Ap
43	37	58.7	766	15	US-10-231-913-99	Sequence 99, Appl
44	37	58.7	766	15	US-10-231-913-101	Sequence 101, App
45	37	58.7	767	15	US-10-262-445-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1

US-10-043-487-340
; Sequence 340, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
; TITLE OF INVENTION: mammalian polypeptides
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2001-01-12
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 340
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Shigella Flexneri
US-10-043-487-340

Query Match 100.0%; Score 63; DB 14; Length 209;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11

Db 45 GDVQCDCIQMV 55

RESULT 2

US-09-978-418-40
; Sequence 40, Application US/09978418
; Publication No. US20030118997A1
; GENERAL INFORMATION:
; APPLICANT: Bejani, Stephan
; APPLICANT: Tanaka, Hiroaki

```

; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 142.U.S.REG
; CURRENT APPLICATION NUMBER: US/09/978,418
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/311,305
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/314,734
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/318,204
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/326,470
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: JPatent
; SEQ ID NO 40
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-418-40

Query Match      100.0%; Score 63; DB 10; Length 479;
Best Local Similarity 100.0%; Pred.No. 0.034;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GDVQDCIQMV 11
Db      195 GDVQDCIQMV 205

RESULT 3
US-09-767-007A-2
; Sequence 2, Application US/09767007A
; Patent No. US2002007275A1
; GENERAL INFORMATION:
; APPLICANT: John S. O'Brien
; APPLICANT: Yasuo Kishimoto
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
; TITLE OF INVENTION: SAPOSIN C AND NEUROTROPHIC PEPTIDES DERIVED THEREFROM
; FILE REFERENCE: MYELOS.2DC1C1
; CURRENT APPLICATION NUMBER: US/09/767,007A
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 08/958,970
; PRIOR FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 523
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-767-007A-2

Query Match      100.0%; Score 63; DB 9; Length 523;
Best Local Similarity 100.0%; Pred.No. 0.037;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GDVQDCIQMV 11
Db      194 GDVQDCIQMV 204

RESULT 4
US-09-870-759-60
; Sequence 60, Application US/09870759
; Patent No. US20020177551a1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166

```

```
RESULT 7
US-09-870-759-61
; Sequence 61, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-61

Query Match      100.0%; Score 63; DB 9; Length 527;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GDVCQDCIQMV 11
Db      195 GDVCQDCIQMV 205

RESULT 8
US-09-751-708A-61
; Sequence 61, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-751-708A-61

Query Match      100.0%; Score 63; DB 10; Length 527;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GDVCQDCIQMV 11
Db      195 GDVCQDCIQMV 205

RESULT 9
US-10-060-036-73
; Sequence 73, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
```

```
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-036-73

Query Match      100.0%; Score 63; DB 14; Length 527;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GDVCQDCIQMV 11
Db      195 GDVCQDCIQMV 205

RESULT 10
US-09-925-298-644
; Sequence 644, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 644
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-298-644

Query Match      90.5%; Score 57; DB 12; Length 156;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2 DVCQDCIQMV 11
Db      13 DVCQDCIQMV 22

RESULT 11
US-10-102-806-644
; Sequence 644, Application US/10102806
; Publication No. US2003054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 644
; LENGTH: 156
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-644

Query Match          90.5%; Score 57; DB 14; Length 156;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2 DVCQDCIQMV 11
Db  13 DVCQDCIQMV 22

RESULT 12
US-10-205-194-176
; Sequence 176, Application US/10205194
; Publication No. US20030134301A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pincock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018201
; CURRENT APPLICATION NUMBER: US/10/205,194
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 176
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: Prosaposin
US-10-205-194-176

Query Match          74.6%; Score 47; DB 14; Length 554;
Best Local Similarity 70.0%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy  2 DVCQDCIQMV 11
Db  195 DVCQDCMKLV 204

RESULT 13
US-10-276-162-1
; Sequence 1, Application US/10276162
; Publication No. US20030215822A1
; GENERAL INFORMATION:
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: YAO, Monique G.
; APPLICANT: BRUNS, Christopher M.
; APPLICANT: YUE, Henry
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: HAFALIA, April
; APPLICANT: PATTERSON, Chandra
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Daniel B.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: JACKSON, Jennifer L.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: BATRA, Sajeev

Query Match          66.7%; Score 42; DB 11; Length 531;
Best Local Similarity 54.5%; Pred. No. 92;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy  1 GDVCQDCIQMV 11
Db  181 GALCQDCVQV 191

RESULT 14
US-09-833-245-903
; Sequence 903, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 903
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-903

Query Match          66.7%; Score 42; DB 11; Length 531;
Best Local Similarity 54.5%; Pred. No. 92;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy  1 GDVCQDCIQMV 11
Db  181 GALCQDCVQV 191

RESULT 15
```



```
US-10-083-357-929
; Sequence 929, Application US/10083357
; Publication No. US20030054370A1
; GENERAL INFORMATION:
; APPLICANT: Qiandong Zeng et al.
; TITLE OF INVENTION: Systemic Discovery of New Genes
; FILE REFERENCE: 032796-090
; CURRENT APPLICATION NUMBER: US/10/083,357
; CURRENT FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 1346
; SEQ ID NO 929
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-083-357-929

Query Match      65.1%; Score 41; DB 14; Length 60;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 GDVCQDCIQ 9
        |||||
Db      28 GDACSDCFQ 36

Search completed: May 5, 2004, 13:28:24
Job time : 1.16999 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:10:37 ; Search time 0.480532 Seconds
(without alignments)
1181.786 Million cell updates/sec

Title: US-09-743-684A-19
Perfect score: 63
Sequence: 1 GDVQCDCIQMV 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	100.0	11	4	US-09-352-548-19
2	63	100.0	80	2	US-08-584-671-15
3	63	100.0	80	3	US-09-027-376-15
4	63	100.0	80	3	US-09-094-192-15
5	63	100.0	81	4	US-09-352-548-2
6	63	100.0	523	1	US-08-100-247-2
7	63	100.0	523	1	US-08-483-146A-2
8	63	100.0	523	1	US-08-232-513A-3
9	63	100.0	523	1	US-08-484-594A-2
10	63	100.0	523	4	US-09-076-258A-2
11	63	100.0	523	4	US-08-756-031-2
12	63	100.0	524	4	US-09-352-548-1
13	57	90.5	10	4	US-09-352-548-21
14	57	90.5	16	4	US-09-352-548-48
15	55	87.3	13	4	US-09-352-548-13
16	53	84.1	9	4	US-09-352-548-24
17	53	84.1	11	4	US-09-352-548-32
18	53	84.1	11	4	US-09-352-548-33
19	51	81.0	9	4	US-09-352-548-29
20	48	76.2	8	4	US-09-352-548-25
21	47	74.6	8	4	US-09-352-548-30
22	47	74.6	19	2	US-08-584-671-5
23	47	74.6	19	3	US-09-027-376-5
24	47	74.6	19	3	US-09-094-192-5
25	47	74.6	79	2	US-08-584-671-14
26	47	74.6	79	2	US-08-584-671-16
27	47	74.6	79	3	US-09-027-376-14

28	47	74.6	79	3	US-09-027-376-16	Sequence 16, Appl
29	47	74.6	79	3	US-09-094-192-14	Sequence 14, Appl
30	47	74.6	79	3	US-09-094-192-16	Sequence 16, Appl
31	43	68.3	7	4	US-09-352-548-26	Sequence 26, Appl
32	43	68.3	11	4	US-09-352-548-20	Sequence 20, Appl
33	40	63.5	15	2	US-08-584-671-4	Sequence 4, Appl
34	40	63.5	15	3	US-09-027-376-4	Sequence 4, Appl
35	40	63.5	15	3	US-09-094-192-4	Sequence 4, Appl
36	39	61.9	6	4	US-09-352-548-27	Sequence 27, Appl
37	39	61.9	16	4	US-09-352-548-55	Sequence 55, Appl
38	38	60.3	7	4	US-09-352-548-31	Sequence 31, Appl
39	36	57.1	103	2	US-09-109-266-4	Sequence 4, Appl
40	36	57.1	252	1	US-08-411-777-8	Sequence 8, Appl
41	36	57.1	252	3	US-09-057-088-8	Sequence 8, Appl
42	36	57.1	341	3	US-09-008-465-1	Sequence 1, Appl
43	36	57.1	341	4	US-09-528-959-1	Sequence 1, Appl
44	35	55.6	52	4	US-09-621-976-6386	Sequence 6386, Ap
45	35	55.6	72	3	US-08-338-579A-101	Sequence 101, App

ALIGNMENTS

RESULT 1
US-09-352-548-19
; Sequence 19, Application US/09352548
; Patent No. 6500431
; GENERAL INFORMATION:
; APPLICANT: Gill, Parkash S.
; APPLICANT: Parkash S. Gill, M.D., Inc.
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
; FILE REFERENCE: 017986-000410US
; CURRENT APPLICATION NUMBER: US/09/352,548
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: US 60/092,647
; EARLIER FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:G1-V11
; OTHER INFORMATION: anti-angiogenic polypeptide
US-09-352-548-19

Query Match 100.0%; Score 63; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDVQCDCIQMV 11
Db 1 GDVQCDCIQMV 11

RESULT 2
US-08-584-671-15
; Sequence 15, Application US/08584671
; Patent No. 5910568
; GENERAL INFORMATION:
; APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,
; APPLICANT: CRAMER, PALMER
; TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM
; TITLE OF INVENTION: TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA
; ADDRESSEE: STATE UNIVERSITY
; STREET: 113 TECHNOLOGY CENTER
; CITY: UNIVERSITY PARK
; STATE: PENNSYLVANIA

COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/584,671
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MONAHAN, THOMAS J
REGISTRATION NUMBER: 29835
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 814-865-6277
TELEFAX: 814-865-3591
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 80
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: UNKNOWN
US-08-584-671-15

Query Match 100.0%; Score 63; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11
Db 67 GDVQCDCIQMV 77

RESULT 3

US-09-027-376-15
; Sequence 15, Application US/09027376
; Patent No. 6004586
; GENERAL INFORMATION:
; APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,
; APPLICANT: CRAWER, PALMER
; TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM
; TITLE OF INVENTION: TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE
; TITLE OF INVENTION: TO ENHANCE OR DECREASE POTENTIAL FERTILITY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA
; ADDRESSEE: STATE UNIVERSITY
; STREET: 113 TECHNOLOGY CENTER
; CITY: UNIVERSITY PARK
; STATE: PENNSYLVANIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 16802-7000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: NEC 286
; OPERATING SYSTEM: DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/027,376
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/584,671
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MONAHAN, THOMAS J
; REGISTRATION NUMBER: 29835
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 814-865-6277
; TELEFAX: 814-865-3591

; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: UNKNOWN
US-09-027-376-15

Query Match 100.0%; Score 63; DB 3; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11
Db 67 GDVQCDCIQMV 77

RESULT 4

US-09-094-192-15
; Sequence 15, Application US/09094192
; Patent No. 6103483
; GENERAL INFORMATION:
; APPLICANT: HAMMERSTEDT, ROY H., BARBATO, GUY F.
; APPLICANT: CRAWER, PALMER
; TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM TO EGG SURFACES AND PROC
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA STATE UNIVERSITY
; STREET: 113 TECHNOLOGY CENTER
; CITY: UNIVERSITY PARK
; STATE: PENNSYLVANIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 16802-7000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: NEC 286
; OPERATING SYSTEM: DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/094,192
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MONAHAN, THOMAS J
; REGISTRATION NUMBER: 29835
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 814-865-6277
; TELEFAX: 814-865-3591
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: UNKNOWN
US-09-094-192-15

Query Match 100.0%; Score 63; DB 3; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11
Db 67 GDVQCDCIQMV 77

RESULT 5

US-09-352-548-2
; Sequence 2, Application US/09352548
; Patent No. 6500431
; GENERAL INFORMATION:
; APPLICANT: Gill, Parkash S.
; APPLICANT: Parkash S. Gill, M.D., Inc.

;; TITLE OF INVENTION: No. 6500431a1 Inhibitors of Angiogenesis and Tumor Growth

;; FILE REFERENCE: 017986-000410US

;; CURRENT APPLICATION NUMBER: US/09/352,548

;; CURRENT FILING DATE: 1999-07-12

;; EARLIER APPLICATION NUMBER: US 60/092,647

;; EARLIER FILING DATE: 1998-07-13

;; NUMBER OF SEQ ID NOS: 59

;; SOFTWARE: Patent in Ver. 2.1

;; SEQ ID NO 2

;; LENGTH: 81

;; TYPE: PRT

;; ORGANISM: Homo sapiens

;; FEATURE:

;; OTHER INFORMATION: Saposin B

;; US-09-352-548-2

Query Match 100.0%; Score 63; DB 4; Length 81;

Best Local Similarity 100.0%; Pred. No. 0.0021;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11

Db 1 GDVCQDCIQMV 11

RESULT 6

US-08-100-247-2

;; Sequence 2, Application US/08100247

;; Patent No. 5571787

;; GENERAL INFORMATION:

;; APPLICANT: O'BRIEN, JOHN S.

;; APPLICANT: KISHIMOTO, YASUO

;; TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR

;; NUMBER OF SEQUENCES: 5

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR

;; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR

;; CITY: NEWPORT BEACH

;; STATE: CA

;; COUNTRY: USA

;; ZIP: 92660

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patent in Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/100,247

;; FILING DATE: 19930730

;; CLASSIFICATION: 514

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Israelaesen, Ned A.

;; REGISTRATION NUMBER: 29,655

;; REFERENCE/DOCKET NUMBER: O'BRIEN.002A

;; TELEPHONE: 619-235-8550

;; TELEFAX: 619-235-0176

;; INFORMATION FOR SEQ ID NO: 2:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 523 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide

;; HYPOTHETICAL: NO

;; ANTI-SENSE: NO

;; FRAGMENT TYPE: N-terminal

;; IMMEDIATE SOURCE:

;; CLONE: PROSAPOSIN

;; US-08-100-247-2

Query Match 100.0%; Score 63; DB 1; Length 523;

Best Local Similarity 100.0%; Pred. No. 0.012;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11

Db 194 GDVCQDCIQMV 204

RESULT 7

US-08-483-146A-2

;; Sequence 2, Application US/08483146A

;; Patent No. 5696080

;; GENERAL INFORMATION:

;; APPLICANT: O'Brien, John S.

;; APPLICANT: Kishimoto, Yasuo

;; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS

;; COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED

;; THEREFROM

;; NUMBER OF SEQUENCES: 11

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Knobbe, Martens, Olson and Bear

;; STREET: 620 Newport Center Blvd. 16th Floor

;; CITY: Newport Beach

;; STATE: CA

;; COUNTRY: USA

;; ZIP: 92660

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Diskette

;; COMPUTER: IBM Compatible

;; OPERATING SYSTEM: DOS

;; SOFTWARE: FASTSEQ for Windows Version 2.0

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/483,146A

;; FILING DATE: 07-JUN-1995

;; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER:

;; FILING DATE:

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Israelaesen, Ned A

;; REGISTRATION NUMBER: 29,655

;; REFERENCE/DOCKET NUMBER: MYELOS.002DV1

;; TELEPHONE: 619-235-8550

;; TELEFAX: 619-235-0176

;; TELEX:

;; INFORMATION FOR SEQ ID NO: 2:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 523 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide

;; FRAGMENT TYPE: N-terminal

;; US-08-483-146A-2

Query Match 100.0%; Score 63; DB 1; Length 523;

Best Local Similarity 100.0%; Pred. No. 0.012;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11

Db 194 GDVCQDCIQMV 204

RESULT 8

US-08-232-513A-3

;; Sequence 3, Application US/08232513A

;; Patent No. 5700909

;; GENERAL INFORMATION:

;; APPLICANT: O'Brien, John S.

;; TITLE OF INVENTION: Prosaoposin and Cytokine-Derived Peptides

;; as Therapeutic Agents

;; NUMBER OF SEQUENCES: 20

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Campbell & Flores LLP
;; STREET: 4370 La Jolla Village Drive, Suite 700
;; CITY: San Diego
;; STATE: California
;; COUNTRY: USA
;; ZIP: 92122
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/232,513A
;; FILING DATE: 21-APR-1994
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/100,247
;; FILING DATE: 30-JUL-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Campbell, Cathryn A.
;; REGISTRATION NUMBER: 31,815
;; REFERENCE/DOCKET NUMBER: P-UD 1643
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (619) 535-9001
;; TELEFAX: (619) 535-8949
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 523 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; FEATURE:
;; NAME/KEY: Protein
;; LOCATION: 1..523
;; OTHER INFORMATION: /label= Hum_prosaposin
US-08-232-513A-3

Query Match 100.0%; Score 63; DB 1; Length 523;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDVQCDCIQMV 11
|||
Db 194 GDVQCDCIQMV 204

RESULT 9
US-08-484-594A-2
; Sequence 2, Application US/08484594A
; Patent No. 5714459
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; APPLICANT: Kishimoto, Yasuo
; TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUROTROPHIC PEPTIDES
; TITLE OF INVENTION: DERIVED THEREFROM
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,594A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/100,247
;; FILING DATE: 30-JUL-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Israel, Ned A.
;; REGISTRATION NUMBER: 29,655
;; REFERENCE/DOCKET NUMBER: MYELOS.002DV2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 619-235-8550
;; TELEFAX: 619-235-0176
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 523 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FRAGMENT TYPE: N-terminal
US-08-484-594A-2

Query Match 100.0%; Score 63; DB 1; Length 523;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDVQCDCIQMV 11
|||
Db 194 GDVQCDCIQMV 204

RESULT 10
US-09-076-258A-2
; Sequence 2, Application US/09076258A
; Patent No. 6559124
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; APPLICANT: Kishimoto, Yasuo
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED
; TITLE OF INVENTION: THEREFROM
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Blvd. 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/076,258A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/958,970
; FILING DATE: 28-OCT-97
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/483,146
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: 08/100,247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: MYELOS.2DV1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; TELEX:

;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 523 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FRAGMENT TYPE: N-terminal
US-09-076-258A-2

Query Match 100.0%; Score 63; DB 4; Length 523;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQDCIQMV 11
|||
Db 194 GDVQDCIQMV 204

RESULT 11
US-08-756-031-2
; Sequence 2, Application US/08756031
; Patent No. 6590074
; GENERAL INFORMATION:
; APPLICANT: O'BRIEN, JOHN S.
; APPLICANT: KISHIMOTO, YASUO
; TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/756,031
; FILING DATE: 26-NOV-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/100,247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: O'BRIEN.002A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; IMMEDIATE SOURCE:
; CLONE: PROSAPOSIN
US-08-756-031-2

Query Match 100.0%; Score 63; DB 4; Length 523;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQDCIQMV 11

Db 194 GDVQDCIQMV 204
|||
|||

RESULT 12

US-09-352-548-1
; Sequence 1, Application US/09352548
; Patent No. 6500431
; GENERAL INFORMATION:
; APPLICANT: Gill, Parkash S.
; APPLICANT: Parkash S. Gill, M.D., Inc.
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
; FILE REFERENCE: 017986-000410US
; CURRENT APPLICATION NUMBER: US/09/352,548
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: US 60/092,647
; EARLIER FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: prosaposin
; NAME/KEY: PEPTIDE
; LOCATION: (195)..(275)
; OTHER INFORMATION: Saposin B
US-09-352-548-1

Query Match 100.0%; Score 63; DB 4; Length 524;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQDCIQMV 11
|||
Db 195 GDVQDCIQMV 205

RESULT 13

US-09-352-548-21
; Sequence 21, Application US/09352548
; Patent No. 6500431
; GENERAL INFORMATION:
; APPLICANT: Gill, Parkash S.
; APPLICANT: Parkash S. Gill, M.D., Inc.
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
; FILE REFERENCE: 017986-000410US
; CURRENT APPLICATION NUMBER: US/09/352,548
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: US 60/092,647
; EARLIER FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:D2-V11
; OTHER INFORMATION: anti-angiogenic polypeptide
US-09-352-548-21

Query Match 90.5%; Score 57; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DVCQDCIQMV 11
|||
Db 1 DVCQDCIQMV 10

RESULT 14
US-09-352-548-48
; Sequence 48, Application US/09352548
; Patent No. 6500431
; GENERAL INFORMATION:
; APPLICANT: Gill, Parkash S.
; APPLICANT: Parkash S. Gill, M.D., Inc.
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
; FILE REFERENCE: 017986-000410US
; CURRENT APPLICATION NUMBER: US/09/352,548
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: US 60/092,647
; EARLIER FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:anti-angiogenic
; OTHER INFORMATION: polypeptide
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (1)-(6)
; OTHER INFORMATION: Xaa = any amino acid, Xaa at positions 1-6 may be
; OTHER INFORMATION: present or absent
US-09-352-548-48

Query Match 90.5%; Score 57; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DVCQDCIQMV 11
Db 7 DVCQDCIQMV 16

RESULT 15
US-09-352-548-13
; Sequence 13, Application US/09352548
; Patent No. 6500431
; GENERAL INFORMATION:
; APPLICANT: Gill, Parkash S.
; APPLICANT: Parkash S. Gill, M.D., Inc.
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
; FILE REFERENCE: 017986-000410US
; CURRENT APPLICATION NUMBER: US/09/352,548
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: US 60/092,647
; EARLIER FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:anti-angiogenic
; OTHER INFORMATION: polypeptide
US-09-352-548-13

Query Match 87.3%; Score 55; DB 4; Length 15;
Best Local Similarity 90.0%; Pred. No. 0.0082;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDVCQDCIQM 10
Db 6 GDVCQDCIQV 15

Search completed: May 5, 2004, 13:16:44
Job time : 0.480532 secs

The present sequence is that of a claimed peptide, derived from human saposin B, that has antiangiogenic activity. The invention is based on the discovery that saposin B (see AAV58716), previously known to be involved in the hydrolysis of sphingolipids, has potent antiangiogenic and antitumour activity, and also has antiproliferative and antimigratory activity against endothelial cells. This activity is conserved in cryptic polypeptides as small as 5 amino acids (see AAV58684-715), which can be synthetically prepared and used *in vitro* or *in vivo* for the treatment of undesired angiogenesis and tumour growth, especially Kaposi's sarcoma (claimed). The polypeptides can also be used in conjunction with cytotoxic moieties to selectively kill certain cell types, e.g. for treatment of cancer, angiofibroma, neovascular glaucoma, arteriovenous malformation, nonunion fracture, arthritis and other connective tissue disorders, Osler-Weber syndrome, atherosclerotic plaque, psoriasis, and other diseases.

CC corneal graft neovascularization, pyogenic granuloma, retrolental
CC fibroplasia, diabetic retinopathy, scleroderma, haemangioma, trachoma,
CC vascular adhesions and hypertrophic scars
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 63; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00095;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMW 11
| | | | | | | | | | |
Db 1 GDVQCDCIQMW 11

RESULT 2
AAB31929
ID AAB31929 standard; protein; 83 AA.
XX
AC AAB31929;
XX
DT 15-MAY-2001 (first entry)
XX
DE Amino acid sequence of a human saposin B protein.
XX
KW Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;
KW ganglioside GM2 activator; saposin B; degenerative disease; glial cell;
KW neurological disease; auto-immune disease; multiple sclerosis; toxicity;
KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
KW rheumatoid polyarthrititis; lupus erythematosus; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200105422-A2.
XX
PD 25-JAN-2001.
XX
PF 17-JUL-2000; 2000WO-FR002057.
XX
PR 15-JUL-1999; 99FR-00009372.
XX
PA (INMR) BIOMERIEUX STELHYS.
XX
PI Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;
XX WPI; 2001-159475/16.
XX
DR N-PSDB; AAF54720.
XX
PT Detecting, preventing and treating degenerative, neurological and
PT autoimmune diseases, particularly multiple sclerosis, using specified
PT polypeptides or related nucleic acid or ligand.
XX
PS Disclosure; Fig 3; 209pp; French.
XX

CC The present sequence represents a human polypeptide, which is used in the
CC method of the invention. The specification describes a method which uses
CC at least one polypeptide or polynucleotide sequence belonging to the
CC perlecan, precursor of the retinol-binding plasma protein, precursor of
CC the ganglioside GM2 activator, calgranulin B or saposin B protein
CC families. The method is used for detecting, preventing or treating a
CC degenerative, neurological and/or auto-immune disease. The
CC polynucleotides and polypeptides are used for diagnosis, prognosis,
CC prevention and treatment of multiple sclerosis (in its various forms and
CC phases). They may also be useful in cases of e.g. Alzheimer's and
CC Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid
CC polyarthrititis and lupus erythematosus, including use as vaccines and in
CC gene therapy (expression of sense or antisense sequences). They can also
CC be used to assess efficacy of potential therapeutic agents, particularly
CC compounds that reduce or inhibit toxicity towards glial cells
XX
SQ Sequence 83 AA;

Query Match 100.0%; Score 63; DB 4; Length 83;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMW 11
| | | | | | | | | | |
Db 3 GDVQCDCIQMW 13

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMW 11
| | | | | | | | | | |
Db 1 GDVQCDCIQMW 11

RESULT 3
AAB31912
ID AAB31912 standard; protein; 85 AA.
XX
AC AAB31912;
XX
DT 15-MAY-2001 (first entry)
XX
DE Amino acid sequence of a human saposin B protein.
XX
KW Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;
KW ganglioside GM2 activator; saposin B; degenerative disease; glial cell;
KW neurological disease; auto-immune disease; multiple sclerosis; toxicity;
KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
KW rheumatoid polyarthrititis; lupus erythematosus; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200105422-A2.
XX
PD 25-JAN-2001.
XX
PF 17-JUL-2000; 2000WO-FR002057.
XX
PR 15-JUL-1999; 99FR-00009372.
XX
PA (INMR) BIOMERIEUX STELHYS.
XX
PI Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;
XX WPI; 2001-159475/16.
XX
DR N-PSDB; AAF54720.
XX
PT Detecting, preventing and treating degenerative, neurological and
PT autoimmune diseases, particularly multiple sclerosis, using specified
PT polypeptides or related nucleic acid or ligand.
XX
PS Claim 1; Page 169; 209pp; French.
XX

CC The present sequence represents a human polypeptide, which is used in the
CC method of the invention. The specification describes a method which uses
CC at least one polypeptide or polynucleotide sequence belonging to the
CC perlecan, precursor of the retinol-binding plasma protein, precursor of
CC the ganglioside GM2 activator, calgranulin B or saposin B protein
CC families. The method is used for detecting, preventing or treating a
CC degenerative, neurological and/or auto-immune disease. The
CC polynucleotides and polypeptides are used for diagnosis, prognosis,
CC prevention and treatment of multiple sclerosis (in its various forms and
CC phases). They may also be useful in cases of e.g. Alzheimer's and
CC Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid
CC polyarthrititis and lupus erythematosus, including use as vaccines and in
CC gene therapy (expression of sense or antisense sequences). They can also
CC be used to assess efficacy of potential therapeutic agents, particularly
CC compounds that reduce or inhibit toxicity towards glial cells
XX
SQ Sequence 85 AA;

Query Match 100.0%; Score 63; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMW 11
| | | | | | | | | | |
Db 3 GDVQCDCIQMW 13

```

RESULT 4
ABU70504
XX ABU70504 standard; protein; 153 AA.
XX
AC ABU70504;
XX
XX 10-JUN-2003 (first entry)
XX
XX Human adipocyte Selected Interacting domain, SID, #135.
XX
XX Human; prey; adipocyte; SID; selected interacting domain; anorectic;
XX antidiabetic; protein-protein interaction; diabetes;
XX yeast 2-hybrid assay; metabolic disorder; obesity.
XX
XX Homo sapiens.
XX
XX WO200286122-A2.
XX
XX 31-OCT-2002.
XX
XX 14-MAR-2002; 2002WO-EP003768.
XX
XX 14-MAR-2001; 2001US-0275734P.
XX
XX (HYBR-) HYBRIGENICS.
XX
XX Legrain P, Daviet L;
XX
XX MPI; 2003-103412/09.
XX N-PSDB; ACAS7048.
XX
XX New complex between two interacting proteins in adipocyte cells, useful
XX for identifying selected interacting domains that modulate protein
XX interactions, or for preventing or treating metabolic disorders such as
XX obesity or diabetes.
XX
XX Claim 6; Page 152; 382pp; English.
XX
XX The invention relates to a complex between two interacting proteins in
XX adipocyte cells, given in the specification. The proteins are identified
XX by selecting a bait protein from a known adipocyte marker and then
XX performing a yeast 2-hybrid selection to isolate prey proteins encoded by
XX members of an adipocyte cDNA library. The proteins are designated SID
XX (RTM) (selected interacting domains) proteins. Also included are a
XX polynucleotide encoding a polypeptide in the adipocyte cells, a
XX recombinant host cell expressing at least one of the interacting
XX polypeptides of the complex, selecting a modulating compound in adipocyte
XX cells, a SID (RTM) polypeptide comprising any of the 738 amino acid
XX sequences given in the specification (including its fragment or variant),
XX a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
XX given in the specification (including its fragment or variant), a vector
XX comprising the SID (RTM) polynucleotide, a recombinant host cell
XX comprising the vector, a protein chip comprising the polypeptides and a
XX record comprising all or part of the data, listed in the specification.
XX The complex, polypeptides, polynucleotides and compounds are useful for
XX preventing or treating metabolic disorders such as obesity or diabetes.
XX The polynucleotides are useful as probes or primers. The complex is
XX particularly useful for identifying selected interacting domains (SID
XX (RTM)) for screening drugs that modulate the protein interaction, thus
XX exhibiting the therapeutic effect. The present sequence represents a SID
XX (prey) protein of the invention
XX
XX Sequence 153 AA;
XX
XX Query Match 100.0%; Score 63; DB 6; Length 153;
XX Best Local Similarity 100.0%; Pred. No. 0.01;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GDVQDCIQMW 11
XX |||||
XX Db 70 GDVQDCIQMW 80

```

```

RESULT 5
ABU70799
XX ABU70799 standard; protein; 153 AA.
XX
XX AC ABU70799;
XX
XX 10-JUN-2003 (first entry)
XX
XX Human adipocyte Selected Interacting domain, SID, #430.
XX
XX Human; prey; adipocyte; SID; selected interacting domain; anorectic;
XX antidiabetic; protein-protein interaction; diabetes;
XX yeast 2-hybrid assay; metabolic disorder; obesity.
XX
XX Homo sapiens.
XX
XX WO200286122-A2.
XX
XX 31-OCT-2002.
XX
XX 14-MAR-2002; 2002WO-EP003768.
XX
XX 14-MAR-2001; 2001US-0275734P.
XX
XX (HYBR-) HYBRIGENICS.
XX
XX Legrain P, Daviet L;
XX
XX MPI; 2003-103412/09.
XX N-PSDB; ACAS7343.
XX
XX New complex between two interacting proteins in adipocyte cells, useful
XX for identifying selected interacting domains that modulate protein
XX interactions, or for preventing or treating metabolic disorders such as
XX obesity or diabetes.
XX
XX Claim 6; Page 254; 382pp; English.
XX
XX The invention relates to a complex between two interacting proteins in
XX adipocyte cells, given in the specification. The proteins are identified
XX by selecting a bait protein from a known adipocyte marker and then
XX performing a yeast 2-hybrid selection to isolate prey proteins encoded by
XX members of an adipocyte cDNA library. The proteins are designated SID
XX (RTM) (selected interacting domains) proteins. Also included are a
XX polynucleotide encoding a polypeptide in the adipocyte cells, a
XX recombinant host cell expressing at least one of the interacting
XX polypeptides of the complex, selecting a modulating compound in adipocyte
XX cells, a SID (RTM) polypeptide comprising any of the 738 amino acid
XX sequences given in the specification (including its fragment or variant),
XX a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
XX given in the specification (including its fragment or variant), a vector
XX comprising the SID (RTM) polynucleotide, a recombinant host cell
XX comprising the vector, a protein chip comprising the polypeptides and a
XX record comprising all or part of the data, listed in the specification.
XX The complex, polypeptides, polynucleotides and compounds are useful for
XX preventing or treating metabolic disorders such as obesity or diabetes.
XX The polynucleotides are useful as probes or primers. The complex is
XX particularly useful for identifying selected interacting domains (SID
XX (RTM)) for screening drugs that modulate the protein interaction, thus
XX exhibiting the therapeutic effect. The present sequence represents a SID
XX (prey) protein of the invention
XX
XX Sequence 153 AA;
XX
XX Query Match 100.0%; Score 63; DB 6; Length 153;
XX Best Local Similarity 100.0%; Pred. No. 0.01;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GDVQDCIQMW 11
XX |||||
XX Db 74 GDVQDCIQMW 84

```

RESULT 6
 ID ABG70166 standard; protein; 209 AA.
 AC ABG70166;
 DT 21-OCT-2002 (first entry)
 DE Human prey protein for Shigella ipaC #30.
 KW Prey protein; ospB; ospD1; ipaD; ipaC; ipaH9.8; ospG; ospC1; Shigella;
 KW shigellosis; bacillary dysentery; antibacterial; yeast two-hybrid system;
 KW protein-protein interaction; SID; selected interacting domain; human.
 XX
 OS Homo sapiens.
 XX
 PN W0200257303-A2.
 XX
 PD 25-JUL-2002.
 XX
 PF 11-JAN-2002; 2002WO-EP000777.
 XX
 PR 12-JAN-2001; 2001US-0261130P.
 XX
 PA (HYBR-) HYBRIGENICS.
 XX
 PI Legrain P;
 XX
 XX WPI; 2002-599706/64.
 DR N-PSDB; ABS51559.
 XX
 PT New complex of protein-protein interactions between a bait Shigella
 PT flexneri polypeptide and a prey mammalian or human placenta polypeptide
 PT for treating or preventing bacillary dysentery in a mammal or human.
 XX
 PS Claim 7; Page 112-113; 162pp; English.
 XX
 CC The invention relates to a complex of protein-protein interactions
 CC between a Shigella flexneri polypeptide (e.g. ospB, ospD1, ipaB, ipaC,
 CC ipaH9.8, ospG and ospC1) and a mammalian polypeptide defined in the
 CC specification. The complexes are formed using the yeast two-hybrid
 CC system. Also included are (1) a recombinant host cell expressing the
 CC interactions between the Shigella flexneri polypeptide and a mammalian
 CC polypeptide defined in the specification; (2) selecting a modulating
 CC compound that inhibits or activates the protein-protein interactions; (3)
 CC a modulating compound obtained from the method of (2); (4) a SID
 CC (selected interacting domain) polypeptide or its fragment or variant
 CC comprising the human polypeptides appearing as ABG70042-ABG70242; (5) a
 CC SID polynucleotide or its fragment or variant comprising encoding the
 CC above polypeptides a vector comprising (5); (6) a recombinant host cell
 CC containing the vector; and (10) a protein chip comprising Shigella
 CC flexneri polypeptide and a mammalian polypeptide defined in the
 CC specification. A pharmaceutical composition comprising the compound,
 CC polypeptide or polynucleotide is useful for treating or preventing
 CC shigellosis (bacillary dysentery) in a human or mammal. The present
 CC sequence represents a human prey protein isolated by the yeast two-hybrid
 CC assay, forming a complex of the invention with a shigella protein
 XX
 SQ Sequence 209 AA;
 Query Match 100.0%; Score 63; DB 5; Length 209;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GDVQCQCIQMW 11
 Db 45 GDVQCQCIQMW 55
 RESULT 7
 ID ABR41750 standard; protein; 385 AA.
 XX

AC ABR41750;
 XX
 DT 02-JUN-2003 (first entry)
 DE Human DITHP biochemical pathway protein.
 XX
 KW Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
 KW cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
 KW neurological disorder; gastrointestinal disorder; transport disorder;
 KW connective tissue disorder; drug screening; proteome analysis;
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
 KW disease model; toxicological testing; transcript imaging;
 KW biochemical pathway.
 XX
 OS Homo sapiens.
 XX
 PN W0200297031-A2.
 XX
 PD 05-DEC-2002.
 XX
 PF 27-MAR-2002; 2002WO-US010056.
 XX
 PR 28-MAR-2001; 2001US-0279619P.
 PR 29-MAR-2001; 2001US-0280067P.
 PR 29-MAR-2001; 2001US-0280068P.
 PR 16-MAY-2001; 2001US-0291280P.
 PR 17-MAY-2001; 2001US-0291829P.
 PR 17-MAY-2001; 2001US-0291849P.
 PR 19-JUN-2001; 2001US-0299428P.
 PR 20-JUN-2001; 2001US-0299776P.
 PR 20-JUN-2001; 2001US-0300001P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 PI Dfour GE, Hillman JL, Xu JY, Tuason O, Yap PE, Ameshey SR;
 PI Daughtery SC, Dam TC, Liu TP, Nguyen DA, Kleefeld Y, Gerstin EH;
 PI Peralta CH, David WH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
 XX
 DR WPI; 2003-129518/12.
 DR N-PSDB; ACC46687.
 XX
 PT Novel human diagnostic and therapeutic polypeptide useful for identifying
 PT test compound which specifically binds to a polypeptide encoded by human
 PT diagnostic and therapeutic polynucleotide, and to induce antibodies.
 XX
 PS Claim 27; SEQ ID NO 1285; 591pp; English.
 XX
 CC The invention relates to novel human diagnostic and therapeutic
 CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
 CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to
 CC polynucleotide sequences at least 90% identical to the dithp cDNA
 CC sequences of the invention; recombinant vectors, host cells and
 CC transgenic organisms comprising a dithp nucleic acid sequence; the
 CC recombinant production of DITHP proteins; antibodies specific for DITHP
 CC proteins; microarrays comprising dithp nucleic acid sequences; methods
 CC for detecting dithp nucleotide and protein sequences; methods of screening
 CC for compounds which specifically bind a DITHP protein; and methods of
 CC assessing the toxicity of test compounds using a dithp hybridisation
 CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
 CC diagnosis of a wide variety of conditions including cancer and other cell
 CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
 CC viral, fungal or parasitic infections; hormonal disorders; metabolic
 CC disorders; neurological disorders; gastrointestinal disorders; transport
 CC disorders; and connective tissue disorders. They may also be used to
 CC screen for modulators of protein activity or gene expression. DITHP
 CC proteins can additionally be used in analysis of the proteome of a tissue
 CC or cell type and to induce antibodies. The dithp nucleic acids are
 CC additionally useful in somatic or germline gene therapy of the disorders
 CC mentioned above, as a source of antisense sequences, as a source of
 CC probes and primers, in genotyping and identification of individuals, in

CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a DTHP protein which is involved in a
CC biochemical pathway. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX

SQ Sequence 385 AA;

Query Match 100.0%; Score 63; DB 6; Length 385;
Best Local Similarity 100.0%; Pred. No. 0.024;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11

Db 53 GDVCQDCIQMV 63

RESULT 8

ABR39442

ID ABR39442 standard; protein; 479 AA.

XX AC ABR39442;

XX DT 12-JUN-2003 (first entry)

XX DE Human GENSET polypeptide clone name SAP-MU-10.

XX KW GENSET; cytostatic; gene therapy; cancer; transgenic; human.

XX OS Homo sapiens.

XX PN WO2003014151-A2.

XX PD 20-FEB-2003.

XX PF 15-OCT-2001; 2001WO-IB002321.

XX PR 10-AUG-2001; 2001US-0311305P.

XX PR 24-AUG-2001; 2001US-0314734P.

XX PR 07-SEP-2001; 2001US-0318204P.

XX PR 01-OCT-2001; 2001US-0326470P.

XX PA (GSEST) GENSET SA.

XX PI Bejanin S, Tanaka H;

XX DR WPI; 2003-256539/25.

XX DR N-PSDB; ABZ76265.

XX PT New GENSET gene, useful for preparing a composition for treating GENSET-related disorders.

XX PS Claim 2; Page 288-289; 301pp; English.

XX CC The invention relates to isolated GENSET polynucleotides and encoded polypeptides. The GENSET gene is useful for preparing a composition for treating GENSET-related disorders e.g., cancer. Sequences ABR39423-448 represent the novel GENSET polypeptide sequences

XX SQ Sequence 479 AA;

Query Match 100.0%; Score 63; DB 6; Length 479;

Best Local Similarity 100.0%; Pred. No. 0.03;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11

Db 195 GDVCQDCIQMV 205

RESULT 9

AAB31916

ID AAB31916 standard; protein; 523 AA.

XX AC AAB31916;

XX DT 15-MAY-2001 (first entry)

XX DE Amino acid sequence of a human protein.

XX KW Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine; ganglioside GM2 activator; saposin B; degenerative disease; glial cell; neurological disease; auto-immune disease; multiple sclerosis; toxicity; Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis; rheumatoid polyarthritis; lupus erythematosus; gene therapy.

XX OS Homo sapiens.

XX PN WO200105422-A2.

XX PD 25-JAN-2001.

XX PF 17-JUL-2000; 2000WO-FR002057.

XX PR 15-JUL-1999; 99FR-00009372.

XX PA (INMR) BIOMERIEUX STELHYS.

XX PI Roeklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;

XX DR WPI; 2001-159475/16.

XX PT Detecting, preventing and treating degenerative, neurological and autoimmune diseases, particularly multiple sclerosis, using specified polypeptides or related nucleic acid or ligand.

XX PS Claim 1; Page 174-175; 209pp; French.

XX CC The present sequence represents a human protein, which is used in the method of the invention. The specification describes a method which uses at least one polypeptide or polynucleotide sequence belonging to the perlecan, precursor of the retinol-binding plasma protein, precursor of the ganglioside GM2 activator, calgranulin B or saposin B protein families. The method is used for detecting, preventing or treating a degenerative, neurological and/or auto-immune disease. The polynucleotides and polypeptides are used for diagnosis, prognosis, prevention and treatment of multiple sclerosis (in its various forms and phases). They may also be useful in cases of e.g. Alzheimer's and Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid polyarthritis and lupus erythematosus, including use as vaccines and in gene therapy (expression of sense or antisense sequences). They can also be used to assess efficacy of potential therapeutic agents, particularly compounds that reduce or inhibit toxicity towards glial cells

XX SQ Sequence 523 AA;

Query Match 100.0%; Score 63; DB 4; Length 523;

Best Local Similarity 100.0%; Pred. No. 0.032;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11

Db 194 GDVCQDCIQMV 204

RESULT 10

ABU05211

ID ABU05211 standard; protein; 523 AA.

XX AC ABU05211;

XX DT 29-JAN-2003 (first entry)

XX DE Human expressed protein tag (EPT) #1877.

XX XX

inflammatory neurodegenerative disease; toxin-induced liver disease.

XX Homo sapiens.

XX WO9912559-A1.

XX 18-MAR-1999.

XX 09-SEP-1998; 98WO-US019216.

XX 09-SEP-1997; 97US-0058352P.

XX 04-JUN-1998; 98US-0088129P.

XX (REGC) UNIV CALIFORNIA.

XX O'brien JS;

XX WPI; 1999-229139/19.

XX N-PSDB; AAX08488.

XX Use of prosaposin receptor agonist.

XX Claim 7; Fig 2; 90pp; English.

XX Prosaposin is a 70kDa glycoprotein which is proteolytically processed to generate saposins A, B, C and D, all of which are similar to each other and have a similar placement of six cysteines, a glycosylation site and conserved proline residues. Prosaposin, saposin C and prosaposin derived peptides (prosaptides) have therapeutic applications in promoting recovery after toxic, traumatic, myocardial ischaemic, degenerative and inherited lesions to the peripheral and central nervous system.

XX Prosaposin receptor agonists (PRAs) inhibit proinflammatory cytokine-induced apoptosis by activation of the Ser/Thr protein kinase Akt. Akt dissociates complexes of Bcl-2 family members, such as BAD-Bcl-2, releasing Bcl-2 and its family members which inhibit caspases, thereby inhibiting apoptosis. An additional mechanism whereby PRAs inhibit apoptosis is by blocking activation of JNK, a proapoptotic signaling component. Within several minutes after binding to the receptor, PRAs block JNK activation induced by tumor necrosis factor- α (TNF α).

XX The activation of JNK by TNF α is another well known mechanism for TNF α -induced, as well as other proinflammatory cytokine-induced apoptosis. The method can be used for inhibiting apoptosis which is caspase-mediated or induced by a proinflammatory cytokine, for example TNF α or interferon- γ . It can be used for inhibiting apoptosis associated with a disorder such as e.g. rheumatoid arthritis, Crohn's disease, irritable bowel syndrome, asthma, cardiac infarction, congestive heart failure, multiple sclerosis, acute disseminated inflammatory leukoencephalitis, progressive multifocal leukoencephalitis, Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's disease, ischemic heart disease, Guillain-Barre disease, traumatic brain injury, traumatic spinal cord injury, alopecia, AIDS dementia, cerebral malaria, HTLV, neuropathy, inflammatory neurodegenerative disease, and toxin-induced liver disease. This 524 N-terminal peptide of prosaposin also acts as a prosaposin receptor agonist

XX Sequence 524 AA;

Query Match 100.0%; Score 63; DB 2; Length 524;

Best Local Similarity 100.0%; Pred. No. 0.032;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GDVCQDCIQMV 11

Db 195 GDVCQDCIQMV 205

RESULT 13

AA58716

ID AA58716 standard; protein; 524 AA.

XX AA58716;

AC AA58716;

XX 25-APR-2000 (first entry)

DT 25-APR-2000 (first entry)

XX

Human prosaposin.

XX

Prosapoin; saponin B; antiangiogenic; angiogenesis inhibitor;

XX

antitumour; antiproliferative; antimigratory; Kaposi's sarcoma; tumour;

XX

human; therapy.

XX

Homo sapiens.

XX

Location/Qualifiers

XX

Key 195..275

XX

Protein /note= "mature saposin B"

XX

Peptide 195..205

XX

/note= "specifically claimed antiangiogenic peptide of Claim 23"

XX

Peptide 196..200

XX

/note= "specifically claimed antiangiogenic peptide of Claim 4"

XX

WO200002902-A1.

XX

20-JAN-2000.

XX

12-JUL-1999; 99WO-US015772.

XX

13-JUL-1998; 98US-0092647P.

XX

(GILL/) GILL P S.

XX

Gill PS;

XX

WPI; 2000-171128/15.

XX

Saposin B derived peptides, useful as inhibitors of angiogenesis and tumor growth.

XX

Disclosure; Page 18; 78pp; English.

XX

The present sequence is that of human prosaposin, a precursor of saposin B. The invention is based on the discovery that saposin B, previously known to be involved in the hydrolysis of sphingolipids, has potent antiangiogenic and antitumour activity, and also has antiproliferative and antimigratory activity against endothelial cells. This activity is conserved in cryptic polypeptides as small as 5 amino acids (see AAY58684 -715), which can be synthetically prepared and used in vitro or in vivo for the treatment of undesired angiogenesis and tumor growth, especially Kaposi's sarcoma (claimed). The polypeptides can also be used in conjunction with cytotoxic moieties to selectively kill certain cell types, e.g. for treatment of cancer, angiofibroma, neovascular glaucoma, arteriovenous malformation, nonunion fracture, arthritis and other connective tissue disorders, Osler-Weber syndrome, atherosclerotic plaque, psoriasis, corneal graft neovascularization, pyogenic granuloma, retrolental fibroplasia, diabetic retinopathy, scleroderma, haemangioma, trachoma, vascular adhesions and hypertrophic scars

XX

Sequence 524 AA;

XX

Query Match 100.0%; Score 63; DB 3; Length 524;

Best Local Similarity 100.0%; Pred. No. 0.032;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GDVCQDCIQMV 11

Db 195 GDVCQDCIQMV 205

RESULT 14

ABU79099

ID ABU79099 standard; protein; 524 AA.

XX ABU79099;

AC ABU79099;

XX 18-JUN-2003 (first entry)

DT 18-JUN-2003 (first entry)

XX DE Lip-TAA binding protein, Prosaposin.
XX KW Superantigen; SAg; staphylococcal enterotoxin; tumour; cancer; apoptosis;
XX KW gene therapy; mammalian cell receptor; cytosolic;
XX KW tumour associated lipid; anergy; T cell; antigen presenting cell; APC;
XX KW tumouricidal immunocyte; antitumour.
XX OS Unidentified.
XX FN US2002177551-A1.
XX XX 28-NOV-2002.
XX PD 30-MAY-2001; 2001US-00870759.
XX PF 31-MAY-2000; 2000US-0208128P.
XX PR (TERM/) Terman D S.
XX PA Terman DS;
XX PI WPI; 2003-361759/34.
XX DR A mammalian cell receptor, useful in the treatment of cancer by binding
XX PT to tumor associated lipids where the binding induces anergy or apoptosis
XX PT in T cells and antigen presenting cells.
XX PS Disclosure; Page; 167pp; English.
XX XX The invention relates to a mammalian cell receptor, useful in the
XX CC treatment of cancer, which binds to tumour associated lipids and induces
XX CC anergy or apoptosis in the T cells and antigen presenting cells (APCs).
XX CC Also included are a mammalian cell useful in the treatment of cancer
XX CC where the receptor which binds tumour associated lipids and induces
XX CC cellular inactivation or death is deleted or functionally deactivated,
XX CC producing (M1) a tumouricidal immunocyte population in vivo in a mammal
XX CC (by allowing tumour associated lipids to contact immunocytes in which
XX CC receptors for immunosuppressive fatty acids, ceramides, glycolipids,
XX CC sphingolipids, glycosphingolipids, phosphoglycolipids, gangliosides,
XX CC sialylated glycans, lipopeptides and proteoglycolipids are inactivated or
XX CC deleted), a construct useful in the treatment of cancer comprising a
XX CC superantigen (SAg) nucleotide inserted into a virus, a mammalian T cell
XX CC useful in the treatment of cancer (where an adaptor protein which
XX CC inhibits T cell activation by tumour associated antigens is deleted or
XX CC functionally deactivated), a composition useful in the treatment of
XX CC cancer (comprising a lipid raft conjugated to a superantigen), producing
XX CC allowing tumour associated lipids to contact immunocytes, in which
XX CC tumouricidal immunocyte population, and administering the tumouricidally
XX CC activated immunocytes to the host), producing (M3) a tumouricidal APC
XX CC population ex vivo in a mammal (by allowing a tumour associated lipid to
XX CC contact APCs in which receptors for the tumour associated lipids are
XX CC inactivated or deleted to produce a tumouricidally activated population,
XX CC and administering APCs to the host), producing a tumouricidal T cell
XX CC population ex vivo in a mammal (by allowing a tumour associated lipids to
XX CC contact T cells, in which adaptor proteins, which inhibit T cell
XX CC activation by tumour associated antigens, are deleted or functionally
XX CC deactivated to produce a tumouricidal population of T cells, and
XX CC administering the tumouricidally activated T cells to the host, or
XX CC allowing a superantigen-lipid raft to contact T cells ex vivo, and
XX CC administering the tumouricidally activated T cells to the host), treating
XX CC (M5) cancer in a mammal (by administering a lipid binding molecule which
XX CC binds immunosuppressive tumour associated lipids in vivo), producing (M6)
XX CC a tumouricidal T cell population in vivo in a mammal (by allowing a
XX CC tumour associated antigen to contact immunocytes in which adaptor
XX CC proteins which inhibit T cell activation by tumour associated antigens
XX CC are deleted or functionally deactivated) and producing (M7) a
XX CC tumouricidal T cell population ex vivo in a mammal comprising allowing a
XX CC superantigen-lipid raft conjugate to contact immunocytes in vivo. The
XX CC receptors, methods and compositions are useful for treating cancers and
XX CC tumours. Bacterial superantigens are co-administered or administered as

CC fusion constructs with anti-tumour proteins or motifs. The present
CC sequence represents a tumour antigen or a motif identifying a tumour
CC antigen, which can be functionally deactivated in the method of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format from the
CC US patent office website at
CC "seqdata.uspto.gov/sequence.html?DocID=20020177551"
XX XX
XX SQ Sequence 524 AA;
XX
XX Query Match 100.0%; Score 63; DB 6; Length 524;
XX Best Local Similarity 100.0%; Pred. No. 0.032;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 GDVCQDCIQMW 11
XX Db 195 GDVCQDCIQMW 205
XX
XX RESULT 15
XX ABU05200
XX ID ABU05200 standard; protein; 524 AA.
XX AC ABU05200;
XX XX
XX DT 29-JAN-2003 (first entry)
XX XX
XX DE Human expressed protein tag (EPT) #1866.
XX XX
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX KW protease; protease inhibitor; transporter; cytoskeletal protein;
XX KW receptor; transcription factor; cancer; MHC;
XX KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX OS Homo sapiens.
XX XX
XX PN WO200278524-A2.
XX PD 10-OCT-2002.
XX XX
XX PF 28-MAR-2002; 2002WO-US009671.
XX PR 28-MAR-2001; 2001US-0279495P.
XX PR 21-MAY-2001; 2001US-0292544P.
XX PR 08-AUG-2001; 2001US-0310801P.
XX PR 01-OCT-2001; 2001US-0326370P.
XX PR 04-DEC-2001; 2001US-0336780P.
XX PR 20-FEB-2002; 2002US-0358985P.
XX XX
XX PA (ZYCO-) ZYCO INC.
XX XX
XX PI Chicx RM, Tomlinson AJ, Urban RG;
XX XX
XX WPI; 2003-040607/03.
XX XX
XX PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX PT cytoskeletal proteins, receptors or transcription factors), useful for
XX PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX XX leukemia.
XX PS Example 2; SEQ ID NO 1866; 134pp; English.
XX XX
XX CC The invention describes a purified polypeptide, which comprises a
XX CC fragment of a kinase, phosphatase, protease, protease inhibitor,
XX CC transporter, cytoskeletal protein, receptor or transcription factor. The
XX CC polypeptide is useful as an immunogenic composition for eliciting in a
XX CC mammal an immunogenic response directed against any of the purified
XX CC polypeptide. The purified polypeptide, or the antibody that binds to this
XX CC polypeptide, is useful for treating cancer. The polypeptide is also
XX CC useful for identifying compounds that binds to a naturally processed
XX CC class I or class II MHC-binding polypeptide. The polypeptides and
XX CC polynucleotides are particularly useful for treating or preventing

CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 524 AA;

Query Match 100.0%; Score 63; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11

|||||

Db 195 GDVCQDCIQMV 205

Search completed: May 5, 2004, 13:31:04
Job time : 1.53561 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:25:43 ; Search time 52.2507 Seconds
(without alignments)
3164.197 Million cell updates/sec

Title: US-09-743-684A-1
Perfect score: 2789
Sequence: 1 MYALFLIASLLGALAGVL.....NTETAAQNAVEKCRHVMN 524

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1950	69.9	554	11	Q8BFQ1 mus musculus
2	1442	51.7	512	13	Q7SY70 xenopus lae
3	1414	50.7	520	13	Q8UVZ4 brachydanio
4	1343	48.2	522	13	Q9DG82 brachydanio
5	1099.5	39.4	525	11	Q8C1C1 mus musculus
6	786	28.2	449	11	Q8BJV5 mus musculus
7	772	27.7	402	11	Q8C1N0 mus musculus
8	552	19.8	953	5	Q8Y125 drosophila
9	545	19.5	241	4	Q8N7t4 homo sapien
10	517	18.5	121	6	P79254 ovis aries
11	511	18.3	965	5	O15997 bombyx mori
12	432.5	15.5	876	5	Q8IMH4 drosophila
13	349.5	12.5	378	11	Q35489 cavia porce
14	343.5	12.3	458	5	Q95X02 naegleria f
15	343.5	12.3	484	5	Q9BKM1 naegleria f
16	334.5	12.0	374	6	Q9TU81 ovis aries

17	327.5	11.7	441	5	Q9U9P3 drosophila
18	292.5	10.5	370	6	P79333 oryctolagus
19	288	10.3	307	5	Q9BKM2 naegleria f
20	279	10.0	294	5	Q95X03 naegleria f
21	279	10.0	456	5	Q94472 dictyosteli
22	251	9.0	316	5	Q86PA4 drosophila
23	220.5	7.9	429	5	Q18276 caenorhabdi
24	213	7.6	213	10	Q9SCT5 arabisidopsis
25	203.5	7.3	217	10	Q9LZW6 arabisidopsis
26	187.5	6.7	243	6	Q9TT05 ovis aries
27	180	6.5	200	5	Q86JD8 dictyosteli
28	179.5	6.4	240	10	Q9AS89 oryza sativ
29	176.5	6.3	402	5	Q18279 caenorhabdi
30	176	6.3	228	6	Q9BDZ9 ovis aries
31	173.5	6.2	174	10	Q9M614 vitis ripar
32	173.5	6.2	507	10	Q9FRW7 nepenthes a
33	172	6.2	188	5	Q86KA8 dictyosteli
34	171.5	6.1	486	10	Q38934 arabisidopsis
35	171.5	6.1	506	10	O6S390 arabisidopsis
36	165.5	5.9	506	10	Q39311 brassica na
37	163.5	5.9	273	10	Q9LUX5 pyrus pyrif
38	163.5	5.9	513	10	O8VYL3 arabisidopsis
39	162.5	5.8	513	10	Q41713 vigna ungui
40	162	5.8	514	10	Q8L6A9 theobroma c
41	161	5.8	107	10	Q8HOR9 cynara card
42	157	5.6	512	10	O04593 arabisidopsis
43	156.5	5.6	514	10	Q94IA2 glycine max
44	155.5	5.6	205	6	Q9N275 ovis aries
45	155.5	5.6	292	10	Q43407 brassica ol

ALIGNMENTS

RESULT 1

ID	Q8BFQ1	PRELIMINARY;	PRT;	554 AA.
AC	Q8BFQ1;			
DT	01-MAR-2003 (Tremblrel. 23, Created)			
DT	01-MAR-2003 (Tremblrel. 23, Last sequence update)			
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)			
DE	Prosaposin.			
GN	PSAP.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NOD; TISSUE=Kidney, and Thymus;			
RX	MEDLINE=22354683; PubMed=12466851;			
RA	The PANTOM Consortium.			
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;			
RT	"Analysis of the mouse transcriptome based on functional annotation of			
RT	60,770 full-length cDNAs."			
RL	Nature 420:563-573(2002).			
DR	EMBL; AK088369; BAC40308.1; --			
DR	EMBL; AK089998; BAC41035.1; --			
DR	MGI; MGI:97783; Psap.			
DR	GO; GO:0005764; C:lysosome; IEA.			
DR	GO; GO:0006655; P:sphingolipid metabolism; IEA.			
DR	InterPro; IPR003119; Sapa.			
DR	InterPro; IPR007856; SapaB_1.			
DR	InterPro; IPR008138; SapaB_2.			
DR	InterPro; IPR008140; SapaB_sub.			
DR	InterPro; IPR008373; Saposin.			
DR	Pfam; PF02199; Sapa; 2.			
DR	Pfam; PF05184; SapaB_1; 4.			
DR	Pfam; PF03489; SapaB_2; 4.			
DR	PRINTS; PR01797; SAPOSIN.			
DR	ProDom; PD001732; SapaB_sub; 3.			
DR	SMART; SM00162; Sapa; 2.			

```
DR SMART; SM00118; SAPB; 4.
SQ SEQUENCE 554 AA; 61050 MW; PF58DB79C7C0C018 CRC64;

Query Match
Best Local Similarity 69.9%; Score 1950; DB 11; Length 554;
Matches 355; Conservative 78; Mismatches 90; Indels 32; Gaps 2;

Qy 1 MYALFLLASLLGALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHLQTVWVKPTVKS 60
Db 1 MYALALFASILLATALTSFVDPDKTCSGSAVLCDKRDVKTAVDCGAVKHCQOMWSKPTAKS 60

Qy 61 LPDICKDVVTAGDMLKONATEBEELVLEKTCDWLPKPNMSASCKETVDSYLPVILDI 120
Db 61 LPDICKVTVTEAGNLKONATEBEELVLEKTCDEWIHDSLSASCKEVDYSLPVILDM 120

Qy 121 IKGEMSRPGVCSALNLCESLQKHLAEHLNQKLESNKIPELDTEVVAPEMANIPLLLY 180
Db 121 IKGEMSNPGVCSALNLCESLQKHLAEHLNQKLESNKIPELDTEVVAPEMANIPLLLY 180

Qy 181 PQDGRSKFPQKNGDVCQDCIQMVTDTIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Db 180 PQDHPRSQPPKANEDVCQDKMLVSDVQTAVKTNSFFIQGFVDHVKECDRLGPGVSDI 239

Qy 241 CKNYISOYSEIATQMMHMQPKEICALVGFCEVKEPMQTLVPKAVSKNVIPELVE 300
Db 240 CKNYVDQYSEVVCQMLHMQPKEICVLAGFCNEVRKVPKTLVPATETIKNIPALEMMD 299

Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFCKWCKLPKSLSEBCEV 360
Db 300 PVEQNVLQAHNVILCQTCQFVNVKSEILVNNATEELLVKGLSNACALLPDPARTKQEV 359

Qy 361 VDTYGSLSILLESVPVCSMLHLCGSG----- 390
Db 360 VGTGFSLLDIFIHEVNPSSLCGVIGLCAARPELVEALEQAPAIVSALLKEPTPPKQPA 419

Qy 391 -TRLPALTVHTVTPKDGDFCEVCKLVGYLDRLNLEKSTKQEIILAELEKGSFLPDPYQK 449
Db 420 QPKQALPAHPVPPQKNGGFCEVCKLVLYLEHNLKSTKEEILAELEKGSFLPDPYQK 479

Qy 450 QCDQFAEYEPVLEILVNDPFCVCLIKIGACPSAKHKLIGTEKICLWGPSYQNTETA 509
Db 480 QCDQFAEYEPVLEILVNDPFCVCKSGVCSAYKIGVCSAYKLLJLIGTEKICVWGPSYQNMETA 539

Qy 510 AQCNVAECHKRHVN 524
Db 540 ARCNVAECHKRHVN 554

RESULT 2
Q7SY70
ID Q7SY70 PRELIMINARY; PRT; 512 AA.
AC Q7SY70;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
```



```
QY 123 GEMSRP-GEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPILLVYDQGRSKPQ 181
D 127 GAQSTDLASVCTALTLCPELQRLHAETTSERPLTQE-----DANEVMAFFLSNGALSFHP 181
QY 182 QDGRSRPQKPDGVCQDCIQMTDIOATVNTSTFVQALVEHVKEECRDLRGLGMDIC 241
D 182 SQMP-----EGAVCHDCVOLISLQDALESNLTAEVTQN---QCQSMGFLAALC 230
QY 242 KNYISQYSEIAIOMMHMPKEICALVGFCDVEKEMPOTLVPAKVASKNVIIPALELVEP 301
D 231 ENYTHROFVPAKOTLOGLPPEVCRKGGFCERESAHWL-----TRVAADVGVPSLEWEMP 285
QY 302 IKHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSSECOEVV 361
D 286 -RTNELQWLQGTCDVCLNLVQELDKWLVNTNSTEALISHTLERVCTVTPPELVQCCITLV 344
QY 362 DTYGSSILSLLEVSPELVCSMLHLCSGTR-----LPALTVHVTPQKDGFC 410
D 345 DTYSPELVQ-LMSKVTPEKVCETIKLCSGKRARSISRAVATPSLP--VDEENQGSFCQ 401
QY 411 VCKLVGYLDRLNLEKNSKQILAALEKGC--SFLPDPYQKQCD 452
D 402 GCKELIGH-----VFPESGPQEHQAG-HSECLQRLWLPDPDVALCD 440
```

RESULT 7

```
Q8CINO Q8CINO PRELIMINARY; PRT; 402 AA.
AC Q8CINO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical saposin A-type domain/saposin type B containing
DE protein.
GN 2310020A21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
DR EMBL; AK009408; BAC25258.1; -
DR MGD; MGI:1924193; 2310020A21RIK.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SApB_1.
DR InterPro; IPR008138; SApB_2.
DR InterPro; IPR008140; SApB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 1.
DR Pfam; PF05184; SApB_1; 3.
DR Pfam; PF03489; SApB_2; 3.
DR PRINTS; PR01797; SApOSIN.
DR ProDom; PD001732; SApB_sub; 2.
DR SMART; SM00162; SAPA; 1.
DR SMART; SM00118; SApB; 3.
KW Hypothetical protein.
SQ SEQUENCE 402 AA; 44420 MW; E90017CBF4017ED6 CRC64;
```

Query Match 27.7%; Score 772; DB 11; Length 402;
Best Local Similarity 38.1%; Pred. No. 2.4e-51;
Matches 154; Conservative 85; Mismatches 129; Indels 36; Gaps 8;

```
QY 131 VCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPILLVYDQGRSKPQ 190
D 13 VCTALTLCPELQRLHAETTSERPLTQE-----DANEVMAFFLSNGALSFHP 62
QY 191 PKONGDVQDCIQMTDIOATVNTSTFVQALVEHVKEECRDLRGLGMDIC 250
D 63 ---EGAVCHDCVOLISLQDALESNLTAEVTQN---QCQSMGFLAALCENYTHROFV 116
QY 251 IAIOMMHMPKEICALVGFCDVEKEMPOTLVPAKVASKNVIIPALELVEP 310
D 117 PAKOTLOGLPPEVCRKGGFCERESAHWL-----TRVAADVGVPSLEWEMP-RTNELQMQ 170
QY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSSECOEVVDTYGSILS 370
D 171 LGLTCDVCLNLVQELDKWLVNTNSTEALISHTLERVCTVTPPELVQCCITLVTPSELVQ 230
QY 371 ILLEVSPELVCSMLHLCSGTR-----LPALTVHVTPQKDGFCVCKLVGYL 419
D 231 -LMSKVTPEKVCETIKLCSGKRARSISRAVATPSLP--VDEENQGSFCQCKRLGMS 287
QY 420 DRNLEKNSKQILAALEKGCFLPDPYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKI 479
D 288 SQNLDHKSTKRDILNAPKGGCRILPLPYVMQCNRFVAEYEPVLIESLKFMNPTDLCKKM 347
QY 480 GACPSAHKPLLGTEKCIWGPSYMCQNTETATAQCNVAVHCKRHVV 523
D 348 GACHGPKTPLLGTQDCVMGSPFWCKSPAAEMCNALSHCQRLVW 391
```

RESULT 8

```
Q9Y125 Q9Y125 PRELIMINARY; PRT; 953 AA.
AC Q9Y125;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SAP-R protein.
GN SAP-R OR BCDA:GH08312 OR CGI2070.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
```

```

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamas I., Simpson M., Skupeki M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agbayani A., Arcania T.T., Baxter E., Blazef R.G., Butenhoff C.,
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA Houston K.A., Hummasti S.R., Kim E., Li P., Moehrefi M., Pacleb J.M.,
RA Park S., Sequeira A., Sethi H., Snir E., Svirskas R.R., Weinburg T.,
RA Cealnikier S.E.;
RA "Full length Drosophila melanogaster cDNA sequence.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF003775; AAF57097.1; -.
DR EMBL; AF145647; AAD38622.1; -.
DR FlyBase; FBgn000416; Sap-1.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SappB_1.
DR InterPro; IPR008138; SappB_2.
DR InterPro; IPR008140; SappB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA_1.
DR Pfam; PF03489; SappB_2; 7.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SappB_sub; 6.
DR SMART; SM00162; SAPA; 1.
DR SMART; SM00118; SAPP; 7.
SQ SEQUENCE 953 AA; 105962 MW; D6CFPD3E9D1502A8 CRC64;

Query Match
Best Local Similarity 19.8%; Score 552; DB 5; Length 953;
Matches 147; Conservative 107; Mismatches 220; Indels 162; Gaps 19;

QY 1 MVALFLASLLGA-ALAGPVLGKCTRGSAVWQNVKTASDCGAVKCLQTWNK---P 56
Db 6 LLAVLALCCAFGVFAATPLGSSKCTWGPSYCGNFSNKECRATHCIQTWETQKVP 65
QY 57 TVKSLPCDICHVVTAAGDMLKONATEEILVYLEKTCMDLKPKNMSASCKEIVDSYLPV 116
Db 66 VDTDSICTCKDMVTQARDQKSNQTEELKEVFEGSKLPIPIKIQEKIVADDFLE 125
QY 117 ILDIIEKEMSPGVCNALNICS-----LQKH----- 144
Db 126 LVEALASQM--PDQVCSVAGLNSARTIDELYKNGIOAGLDGTQVQNEEDSSSETELAMQP 184
QY 145 -----LAELNHQKQLESNKIPELDMTEV-----APPMANIPILL----- 179
Db 185 QLSGCGNCLLSRLMHSKFAATDRD---DMVETMLHWCGLSISFSDACANIVITYFNDYD 241
QY 180 -----YQDGRSKPQPK-----DNGD-----VQDCICQMTVDIQ 209
Db 242 HVSKHLTDAVCHVSGVCASTRYHQHEEKQPOEALVALDAGDDIPCELCEQLVKHLRDVL 301
QY 210 TAVRTNSTFFQALVEHKECDRLGPGMADICKNYISQYSEAIQ-MMMHMQPKICALV 268
Db 302 VANTTEFEFQVMEGCKQ-----SKGFKDECLISVDQYHYIYETLVSKLDANGACCM 356
QY 269 GFCDE-----VKEMPQTLVPAKASKNVIPALEL---VEPIKKE----- 306

```



```
QY 405 --DGGFCVCKLVGYLDRLNLEKNTKQETLAALAEKCSFLPDPYQKQCFVAEYEPVL 462
DB 122 ENQGSFNGCKRLLTSSHNLESSTKRDILVAFKGGCSILPLPYMIQCKHFVQYEPVL 181
QY 463 IEILVEWDSFVCLKIGACSAHKPLIGTEKCTWGSYWCNTTETAACNAVCHCKRHV 522
DB 182 IESLKMMDPVAVCKVYGGACHGPRPTLLGTDDQALGFSFMCRSQEAALKCNVQHCQKHV 241

RESULT 10
P79254
ID P79254 PRELIMINARY; PRT; 121 AA.
AC P79254;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sulfated glycoprotein-1/SGP-1 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96249303; PubMed=8848570;
RA Spencer T.E., Graf G.H., Bazer F.W.;
RT "Sulfated glycoprotein-1 (SGP-1) expression in ovine endometrium
during the oestrous cycle and early pregnancy.";
RL Reprod. Fertil. Dev. 7:1053-1060(1995).
DR EMBL; S82555; AAD14405.1; -.
DR GO; GO:0005764; C:lysosome; IEA.
DR InterPro; IPR008138; SGPB_2.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF03489; SGPB_2; 1.
DR PRINTS; PR01797; SAPOSIN.
DR SMART; SM00118; SGPB; 1.
FT NON TER 1
SQ SEQUENCE 121 AA; 13604 MW; 4F0F5A6EB83D0C9A CRC64;

Query Match 18.5%; Score 517; DB 6; Length 121;
Best Local Similarity 77.7%; Pred. No. 2.5e-32;
Matches 94; Conservative 18; Mismatches 9; Indels 0; Gaps 0;

QY 205 VTDTQTAVTNTSTFQALVEHVKEECRLGPGMADICKNYISOYSEIAIQMMHMQPKEI 264
DB 1 VTDTQTAVTNTSTFQGLVDHVKEECRLGPGMADICKNYLAQYSEIAVQMMHMQPKEI 60

QY 265 CALVGFCDVEKMPQTLVPAKVASKNVI PALELVEPIKKEHVPAKSDVYCEVCEFLVKE 324
DB 61 SALVGFCDVEKESPARTLDPKAEASKNLPALTEPNKQEI PAQTLIFCQVCQFVRE 120

QY 325 V 325
DB 121 V 121

RESULT 11
O15997
ID O15997 PRELIMINARY; PRT; 965 AA.
AC O15997;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BmPi09.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
OX NCBI_TaxID=7051;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=98278844; PubMed=9611271;
RA Tambunan J., Chang P.-K., Li H., Natori M.;
RT "Molecular cloning of a cDNA encoding a silkworm protein which
contains the conserved BH regions of Bcl-2 family proteins.";
RL Gene 212:287-293(1998).
DR EMBL; AB008449; BAA23126.1; -.
DR PIR; T00207; T00207.
DR GO; GO:0005764; C:lysosome; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SGPB_1.
DR InterPro; IPR008138; SGPB_2.
DR InterPro; IPR008140; SGPB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02139; SAPA; 2.
DR Pfam; PF05184; SGPB_1; 6.
DR Pfam; PF03489; SGPB_2; 7.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SGPB_sub; 6.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SGPB; 7.
SQ SEQUENCE 965 AA; 108825 MW; F41A7BEE7F626078 CRC64;

Query Match 18.3%; Score 511; DB 5; Length 965;
Best Local Similarity 24.6%; Pred. No. 1e-30;
Matches 139; Conservative 104; Mismatches 229; Indels 94; Gaps 18;

QY 19 VLGLKSCTRGSVAVCONVKTASDCGAVKHCLOTVWNK---PTVKSLPCDICKVVTAAAGD 75
DB 170 LLGKSRCTGWSYWCNFSFTGRECNATPHCINRWSKMTFPEDNDNICQICLDMVKOARD 229

QY 76 MLKDNATEREILVLEKTDWLPKPNMSASCKEIVSYLVPVLDI IKGMSRCPCEVCAL 135
DB 230 QLOSNETQDEIKVFEGSCKLIPIKVAECMKLADEFVVELIETLASEMN-POAVCSVA 288

QY 136 NLCS--LQKHLAELNHQKLES-----NKPELDMTEVYVAPFMA----- 173
DB 289 GLCNNAKIDRLLDVYNAQRELGRAGCYNQCTGVWRKKFDTEKYEDFLVGLQVCRNMS 348

QY 174 ---NIPLLLYP-----QDG-----PRSKPQPKDNGD 196
DB 349 LSSDCSMLIPKYYENILEAVKDLNPEGICHVSGQCSYKFNHDEFTFPQMVQYSATDD 408

QY 197 V-QDCICIQMTDICTAVRTNSTFQALVEHVKEECRLGPGMADICKNYISOYSEIAIQM 255
DB 409 VPCEFCQLVKHURDLVAVNTTELE-FYKVLQGLCKQTGK-FKDECIHLAEQYYPVYNF 466

QY 256 MM-HMQPKETCALVGFCDVEKMPQTLV-----PAKVA-----SKNVIP 294
DB 467 LVSDLPKPAETCKMIGICGNLTSAPI SPLVARELVVVKVQPKLIGAESKIARVPLAKQMEP 526

QY 295 ALELVE--PIKKHEVPA-KSDVYCEVCEFLVKVTKLIDNNKTEKEITLDAFKMCKLPK 351
DB 527 ASAAVSVPLERMFVAAPQSKAACAFQCYFLHYLQVLSDTREDKVKAAVQEAACDALPD 586

QY 352 SLSEECQVVDYTGSSITSLILEEVSPELVCSMLHLCSGTRLPALTVHVTPQKDGGEV 411
DB 587 ALNGECKEFVTQYGSVAIVALLVQEI DPASVCPALQICPQTE-EIRRVDMVSEKSN--CPL 643

QY 412 CKKLGVYLDRLNLEKNTKQETLAALAEKCSFLPDPYQKQCFVAEYEPVLIEILVEWMD 471
DB 644 CLFAVQLESVLKNNRSEENIRKALDGLCTRLSKQLQSECIDFDVDTYSSQLVEMLVADMN 703

QY 472 PSFVCLKIGAC-PSAHKPLIGTEKCI 496
DB 704 AKEICVFLKLCRDQLDHPDLKLTHSSI 729

RESULT 12
Q8IMH4 PRELIMINARY; PRT; 876 AA.
ID Q8IMH4
```

AC Q81MH4;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE CG12070-PB.
 DE SAP-R.
 GN Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blaise R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaesner K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimble B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Smith T.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "The genome sequence of Drosophila melanogaster."
 RT Science 287:2185-2195(2000).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacleb J., Paragov S., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RA "Sequencing of Drosophila melanogaster genome."
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RN SEQUENCE FROM N.A.
 RP Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochownik S.E., Smith C.D.,

RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celniker S.E.,
 RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
 RA Harris N.L., Kronmiller B., Marshall B., Millburn G.H., Richter J.,
 RA Russo S., Searle S.M.J., Smith E., Shu S., Smutniak F.,
 RA Whitfield E.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
 RA Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003775; AAN14261.2; -
 DR GO; GO:0005764; C:lysosome; IEA.
 DR GO; GO:0005489; F:electron transporter activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006665; P:phospholipid metabolism; IEA.
 DR InterPro; IPR000345; CytC heme_BS.
 DR InterPro; IPR007856; SapB_1.
 DR InterPro; IPR008138; SapB_2.
 DR InterPro; IPR008140; SapB_sub.
 DR InterPro; IPR008373; Saposin.
 DR InterPro; IPR008139; SaposinB.
 DR Pfam; PF05184; SapB_1; 3.
 DR Pfam; PF03489; SapB_2; 4.
 DR PRINTS; PR01797; SAPOSIN.
 DR ProDom; PD001732; SapB_sub; 2.
 DR SMART; SM00118; SAPB; 7.
 DR PROSITE; PS00190; CYTOCHROME_C; 1.
 DR SEQUENCE 876 AA; 97617 MW; 767E16D35ACF52DB CRC64;
 SQ
 Query Match 15.5%; Score 432.5; DB 5; Length 876;
 Best Local Similarity 23.8%; Pred No. le-24;
 Matches 117; Conservative 91; Mismatches 178; Indels 105; Gaps 17;
 QY 61 LPDIDCKDVVTAAGDMLKDNATEBEILVLEKTDMLPKNMSASCKEIVDSYLPVLDI 120
 DB 208 IPCEICQVLKRLDLVANTTETEFKQVMEGCK--QSKGFKDECLISVDQYHYIYET 265
 QY 121 IKGESRRGEVCSALNLC-----MQLPIDHLMGAANPGALVEGSELCTLCYMLHFI 151
 DB 266 LVSKLDANG-ACCIGICQKNASMKDVPIMPLLPVIEPAQVKITIEKLEK-----EK 319
 QY 152 KQLESNKIPELDMTVEVWAPFMANIFLLLYPDQDGRSKPQP---KNGDVGQDCIQWVTDI 208
 DB 320 KQLGASE-PKFSQOEILD-----MQLPIDHLMGAANPGALVEGSELCTLCYMLHFI 370
 QY 209 QTAVRTNSTFVOALVEHVKEE-CDRLGPGMADICKNYSIQTSEIATQMMH-MQPKIECA 265
 DB 371 QETLATSTDD--IKHTVENICAKLPESGVAGQCRNEVEMYGDAVIALVQGLNPRDVP 428
 QY 257 LVGFCDEVKEMPOTLVPAKVASQNVIPALELVEPIKKEHVPAKSDVYCEVCFYKEVT 326
 DB 429 LMQCP--KNLPK-----EDVEFNPPQASDEQDP-----PT-----CPLCLFAVEQAQ 471
 QY 327 KLINNKTEKILDAFKMKSKLPKSLSEECQWVDYTGSSILSILEEVSPELVCSMLH 386
 DB 472 MKIRDNKSQKNIKKVNLGLCSLHPNEIKBECVDFVNTYSNELIDMLITDPKPOEICVOLK 531
 QY 387 LCSGT--RLPALTVHVTQPKDG-----DDVDGDKSSSEISFNDSIESLELPPQAFDPG 411
 DB 532 LCPKTTVALWDLRLISLEDDVDGDKSSSEISFNDSIESLELPPQAFDPGFTAAPNCLI 591
 QY 412 CKKLGVGLDRLNLEKNSTKOBILAALEKGCSEFLPDYKQDQFVAEYEPVLIEILVEVMD 471
 DB 592 CEELVKTLEKMGKHPTRDSIKHILESQDRMRKPNWTKCHKVIDKYGDKIADLLKEND 651
 QY 472 PSFVCLKIGAC 482

```
Db 652 PKLICTELGMC 662
RESULT 13
O35489 PRELIMINARY; PRT; 378 AA.
AC O35489
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Surfactant protein-B.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RA Bingle C.D., Yuan H.T., Gowan S.;
RT "Guinea pig surfactant protein-B is differentially polyadenylated.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF033190; AAB87086.1; -.
DR HSSP; P07988; 1DFW.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0007585; P:respiratory gaseous exchange; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; SApA.
DR InterPro; IPR007856; SApB.1.
DR InterPro; IPR008138; SApB.2.
DR InterPro; IPR008140; SApB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR InterPro; IPR008137; Surfactant_B.
DR Pfam; PF02199; SAPA; 1.
DR Pfam; PF05184; SApB.1; 2.
DR Pfam; PF03489; SApB.2; 3.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SApB_sub; 1.
DR ProDom; PD008002; Surfactant_B; 1.
DR SMART; SM00162; SAPA; 1.
DR SMART; SM00118; SApB; 3.
SQ SEQUENCE 378 AA; 41677 MW; E3DA2E237ED401DA CRC64;

Query Match 12.5%; Score 349.5; DB 11; Length 378;
Best Local Similarity 22.9%; Pred. No. 9.2e-19;
Matches 113; Conservative 67; Mismatches 175; Indels 139; Gaps 15;

Qy 4 LFLASLIGALAGVLGLKECTRGSAVVCQVKTASDCGAVKCIQTVWVKPTVKSIPC 63
Db LVLLPMLCGPAAQAQTTSSSLACSQGPKFCWCSLEQALQCRALGHCLQEVWGHVGADDL-C 68

Qy 64 DICKDVVTAAGDMLKDNATEEELVLEKTCDWLPKPNMSCKEIVDSYLPVLIDTIKG 123
Db QECEDIVRLITMAKEVIFQKIRFLERECDVLPKLLVPRCHSVLETYPFLVIDYFQS 128

Qy 124 EMSRPGVCSALNTCESLQKHLAEINLHOKLESNKIPELDMTEVVAPPMANIPLLLYPD 183
Db HIT-PKTCWNLGLCQPRQ-----PD----- 148

Qy 184 GPRSKPQKNGDVQCQDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLPGMGADICKN 243
Db -----POPE----- 149

Qy 244 YISQVSEIAIQMMHMQPEICALVGFCEVKEPMQTLVPKAV-ASKNVIIPALELVEPI 302
Db SLSETSP-----DLLVPPRLPRALUSVWPGPH-TQDL 191

Qy 303 KKEVPAKSDVYCEYCEFLVKEVTKLIDNNKTEKILD-AFDKMSCKLPKLSLSECEQEV 361
Db SEQQFPPIPLP-YCRICKTLKRVQAMI-----PKGVLAWAQAQCHVVPVLVAGGICQCLA 245

Qy 362 DTYGSSILSILLESVPELVCSMLHLCS-----GTRLPALTVHTVTPQKDGFCFCEVKLV 416
Db ERYTVLLDALLSHLLPOLVGLVLRCSMDNSAGLVLPALSLSEGVLPODSECDLCLSVT 305

Qy 417 GYLDRNLKRNKSTKOEILAALEKGCFLPDPYQKQ-CDQFVAEYEPVLIELVEMDFSFV 475
Db 306 -----TRAWNSSELARPOAMHQAC--LSSPLDRQKCKQFVEQYTFQLLALPRGGDPRTT 358

Qy 476 CLKIGACPSAHKPL 489
Db 359 COALGVCRGAKNPL 372

RESULT 14
Q9SX02 PRELIMINARY; PRT; 458 AA.
AC Q9SX02;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Naegleria fowleri.
GN NP-B.
OS Naegleria fowleri.
OC Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.
OX NCBI_TaxID=5763;
RN [1]
RP SEQUENCE FROM N.A.
RA Nickel R., Leippe M.;
RT "Pore-forming peptides of Naegleria fowleri.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF196309; AAL01158.1; -.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR007856; SApB.1.
DR InterPro; IPR008138; SApB.2.
DR InterPro; IPR008140; SApB_sub.
DR InterPro; IPR008173; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF05184; SApB.1; 5.
DR Pfam; PF03489; SApB.2; 5.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SApB_sub; 4.
DR SMART; SM00118; SApB; 5.
FT NON TER 1
SQ SEQUENCE 458 AA; 50452 MW; 433F60CCD6E4C904 CRC64;

Query Match 12.3%; Score 343.5; DB 5; Length 458;
Best Local Similarity 25.6%; Pred. No. 3.4e-18;
Matches 115; Conservative 69; Mismatches 186; Indels 79; Gaps 16;

Qy 42 CGAVKHC---LQTVWVKPTVKSIPCDICKDVVTAAGDMLKDNATEEELVLEKTCDWLP 98
Db CSQVKICNGSAMAAVAPKAENSGICNNMCQLLVTOQENWVESNDTMTLEKLEQVCSVIP 128

Qy 99 KPNMSCKEIVDSYLPVLIDTIKEMSRPG-EVCSALNLCESLQKHLAEINLHOKLESN 157
Db -GOVSALCTVAVBOYLPFIHQV--EKQFPALTICQDVHLCSAQ----- 170

Qy 158 KIPELDMTEVVAPPMANIPLLLYPDQDGRSKPQKNGDVQCQDCIQMVTDIQTAVRTNST 217
Db -----APVV-----QQQAAELCPICKAAVGLTKI--NNV 201

Qy 218 FVQALVHVKEECDRLPGMGADICKNVIQSYSEIAIQMMHMQPEICALVGFCEVKEKEM 277
Db DVAVAKQLEFACSF--QVPDCCQIVDRAAQA--QDLQTEDAQTCSTV-----VDVC 252

Qy 278 PMQTLVPAKVASKNVIIPALELVEPIKKEHVPAKSDVYCEYCEFLVKEVTKLIDNNKTEK 337
Db PKQQW-----TFNPFKKF-LEAKDSKYCPTCLQITKYLEDLIVSDITVNE 297

Qy 338 ILDAFDKMSCKLPKLSLSECEQEVVDITYGSSILSILLESVPELVCSMLHLCSGTRLPALT 397
Db IIKLADAGCARL-GALESCKKFPVLAVDELKLLLEKLPQKVCSTLKMCDAAELKLA 356

Qy 398 VHTVTPQKDGFCFCEVKLVGLVLRNLEKNSTKQSBILAALKEGC-SFLPDPYQKQCDQFVA 456
Db 398 VHTVTPQKDGFCFCEVKLVGLVLRNLEKNSTKQSBILAALKEGC-SFLPDPYQKQCDQFVA 456
```

```

357 L-APQAAQDTMCLACEYVISVADNMLIANNTQOSVKNTLDKVCQEFVPSIQSQCIALVN 415
457 EYEFVLIEIL-VEVMDPSPVCLKIGACPS 484
    ||| |||:: :||| |||
416 QYEALQILFESKVFNPQTCKAIGVCS 444

RESULT 15
Q9BKM1 PRELIMINARY; PRT; 484 AA.
ID AC
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Naegleriaflore B pore-forming peptide.
OS PRONP-B.
GN Naegleria fowleri.
OC Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.
OX NCBI_TaxID=5763;
RN [1]
RP SEQUENCE FROM N.A.
RA Nickel R., Benkert C., Jacobs T., Marti T., Marciano-Cabral F.,
RT Leippe M.;
RL "pore-forming peptides of Naegleria fowleri.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF154047; AAK21659.1; -
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF05184; SapB_1; 5.
DR Pfam; PF03489; SapB_2; 5.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 4.
DR SMART; SM00118; SAPB; 5.
DR SEQUENCE 484 AA; 53237 MW; 63DA3AAE7CA578B0 CRC64;

```

```

Db          :|||:::|:|:|:|:|
442 QYEAQLIQLFESKVENPQTCKAIGVSS 470

Search completed: May 5, 2004, 13:33:40
Job time : 54.2507 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:16:48 ; Search time 12.9383 Seconds
(without alignments)
2108.841 Million cell updates/sec

Title: US-09-743-684A-1

Perfect score: 2789

Sequence: 1 MYALFLILASLLGAALAGPVL.....NTETAAQCNAVERKRRHVMN 524

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	2789	100.0	524	1	SAP_HUMAN
2	2425.5	87.0	525	1	SAP_BOVIN
3	1996	71.6	554	1	SAP_RAT
4	1938.5	69.5	557	1	SAP_MOUSE
5	1679.5	60.2	518	1	SAP_CHICK
6	393	14.1	80	1	SAP_PIG
7	382	13.7	377	1	PSPB_MOUSE
8	379	13.6	376	1	PSPB_RAT
9	349.5	12.5	381	1	PSPB_HUMAN
10	310	11.1	363	1	PSPB_CANFA
11	296.5	10.6	370	1	PSPB_RABIT
12	261	9.4	81	1	SAP_CAVPO
13	173.5	6.2	513	1	ASPR_CUCPE
14	158.5	5.7	508	1	ASPR_HORVU
15	154.5	5.5	473	1	CYPI_CYNCA
16	154	5.5	496	1	ASPR_ORYSA
17	143.5	5.1	1175	1	GLGI_MOUSE
18	142.5	5.1	509	1	APRI_ORYSA
19	138.5	5.0	1171	1	GLGI_RAT
20	138.5	5.0	1179	1	GLGI_HUMAN
21	137.5	4.9	1160	1	GLGI_CRIGR
22	137.5	4.6	1142	1	GLGI_CHICK
23	120	4.3	578	1	EZRA_ENTFA
24	119	4.3	975	1	SECB_MOUSE
25	118	4.2	857	1	I2C1_HUMAN
26	117	4.2	857	1	I2C1_MOUSE
27	116.5	4.2	1216	1	P1B1_MOUSE
28	116	4.2	975	1	SECB_RAT
29	114.5	4.1	971	1	RECK_HUMAN
30	114.5	4.1	1216	1	P1B1_RAT
31	114	4.1	984	1	NOP_DROME
32	113.5	4.1	974	1	SECB_HUMAN
33	113	4.1	418	1	TEK1_HUMAN

RESULT 1

ID	SAP_HUMAN	STANDARD;	PRT;	524 AA.
AC	P07602;	P15793;	P78538;	P78541; P78546; P78547; P78558;
AC	Q92739;	Q92740;	Q92741;	Q92742;
DT	01-APR-1988	(Rel. 07, Created)		
DT	01-APR-1990	(Rel. 14, Last annotation update)		
DE	Proactivator polypeptide precursor (Contains: Saposin A (Protein A); Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside sulfate activator) (CSAct) (Dispersin) (Sulfatide/GMI activator); Saposin C (Co-beta-glucosidase) (Al activator) (Glucosylceramidase activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D (Protein C) (Component C)).			
GN	PSAP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=90129043; PubMed=25151510;			
RA	Rorman E.G., Grabowski G.A.;			
RT	"Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphingolipid hydrolase activator proteins are encoded by single genes in humans and rats.";			
RL	Genomics 5:486-492(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89255151; PubMed=2498298;			
RA	Nakano T., Sandhoff K., Stuenkel J., Christomanou H., Suzuki K.;			
RT	"Structure of full-length cDNA coding for sulfatide activator, a Co-beta-glucosidase and two other homologous proteins: two alternate forms of the sulfatide activator.";			
RL	J. Biochem. 105:152-154(1989).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM SAP-MU-0).			
RC	TISSUE=Brain, Eye, and Skin;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,			

ALIGNMENTS

34	113	4.1	1102	1	MYSC_CHICK
35	113	4.1	2869	1	RBPI_PLAVB
36	112	4.0	1184	1	PBL2_HUMAN
37	112	4.0	1216	1	P1B1_BOVIN
38	111.5	4.0	8545	1	ANC1_CAEEL
39	110.5	4.0	1971	1	MC3A_MOUSE
40	110.5	4.0	3911	1	AKA9_HUMAN
41	110	3.9	476	1	VTDB_RABIT
42	110	3.9	1216	1	P1B1_HUMAN
43	109	3.9	861	1	I2C4_HUMAN
44	109	3.9	1557	1	DVAL_DICVI
45	108.5	3.9	3674	1	SPCR_HUMAN

P29616	gallus gall
Q00798	plasmodium
P98095	homo sapien
P10894	bos taurus
Q9n4n4	caenorhabdi
Q9wud9	mus musculu
Q99996	h a-kinase
P53789	oryctolagus
Q9nq66	homo sapien
Q9hck5	homo sapien
Q24702	dictyocaulu
Q9nrc6	homo sapien

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences."
 RP Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RA [4]
 RP SEQUENCE OF 59-125 AND 304-513 FROM N.A.
 RA TISSUE=Brain;
 RX MEDLINE=91192146; PubMed=2013321;
 RA Holtschmidt H., Sandhoff K., Fuerst W., Kwon H.Y., Schnabel D.,
 RA Suzuki K.;
 RT "The organization of the gene for the human cerebroside sulfate
 RT activator protein."
 RL FEBS Lett. 280:267-270(1991).
 RN [5]
 RP SEQUENCE OF 164-524 FROM N.A.
 RX MEDLINE=88068647; PubMed=2825202;
 RA Dewji N.N., Wenger D.A., O'Brien J.S.;
 RT "Nucleotide sequence of cloned cDNA for human sphingolipid activator
 RT protein 1 precursor."
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8652-8656(1987).
 RN [6]
 RP PARTIAL SEQUENCE OF 60-142.
 RX MEDLINE=89240739; PubMed=2717620;
 RA Morimoto S., Martin B.M., Yamamoto Y., Kretz K.A., O'Brien J.S.,
 RA Kishimoto Y.;
 RT "Saposin A: second cerebroside activator protein."
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3389-3393(1989).
 RN [7]
 RP SEQUENCE OF 195-263 FROM N.A.
 RX MEDLINE=86130593; PubMed=2868718;
 RA Dewji N.N., Wenger D.A., Fujibayashi S., Donoviel M., Esch F.,
 RA Hill F., O'Brien J.S.;
 RT "Molecular cloning of the sphingolipid activator protein-1 (SAP-1),
 RT the sulfate sulfatase activator."
 RL Biochem. Biophys. Res. Commun. 134:989-994(1986).
 RN [8]
 RP SEQUENCE OF 195-274.
 RC TISSUE=Kidney;
 RX MEDLINE=91006165; PubMed=2209618;
 RA Furst W., Schubert J., Machleidt W., Meyer H.E., Sandhoff K.;
 RT "The complete amino-acid sequences of human ganglioside GM2 activator
 RT protein and cerebroside sulfate activator protein."
 RL Eur. J. Biochem. 192:709-714(1990).
 RN [9]
 RP SEQUENCE OF 195-274.
 RX MEDLINE=89207118; PubMed=3242555;
 RA Kleinschmidt T., Christomanou H., Braunitzer G.;
 RT "Complete amino-acid sequence of the naturally occurring A2 activator
 RT protein for enzymic sphingomyelin degradation: identity to the
 RL sulfate activator protein (SAP-1)."
 RN Biol. Chem. Hoppe-Seyler 369:1361-1365(1988).
 RP SEQUENCE OF 311-390.
 RX MEDLINE=88163077; PubMed=3442600;
 RA Kleinschmidt T., Christomanou H., Braunitzer G.;
 RT "Complete amino-acid sequence and carbohydrate content of the
 RT naturally occurring glucosylceramide activator protein (A1 activator)
 RT absent from a new human Gaucher disease variant."
 RL Biol. Chem. Hoppe-Seyler 368:1571-1578(1987).
 RN [11]
 RP SEQUENCE OF 407-484.
 RX MEDLINE=89000190; PubMed=3048308;
 RA Furst W., Machleidt W., Sandhoff K.;
 RT "The precursor of sulfate activator protein is processed to three
 RT different proteins."
 RL Biol. Chem. Hoppe-Seyler 369:317-328(1988).
 RN [12]
 RP PARTIAL SEQUENCE OF 405-484.
 RX MEDLINE=89025876; PubMed=2845979;
 RA Morimoto S., Martin B.M., Kishimoto Y., O'Brien J.S.;
 RT "Saposin D: a sphingomyelinase activator."
 RL Biochem. Biophys. Res. Commun. 156:403-410(1988).
 RN [13]

RP SEQUENCE OF 17-26.
 RC TISSUE=Milk;
 RX MEDLINE=92069206; PubMed=1958198;
 RA Kondoh K., Hinenio T., Sano A., Kakimoto Y.;
 RT "Isolation and characterization of prosaposin from human milk."
 RL Biochem. Biophys. Res. Commun. 181:286-292(1991).
 RN [14]
 RP PARTIAL SEQUENCE (SAPOSIN B), AND STRUCTURE OF CARBOHYDRATES.
 RC TISSUE=Urine;
 RX MEDLINE=20032116; PubMed=10562467;
 RA Fluharty A.L., Lombardo C., Louis A., Stevens R.L., Whitelegge J.P.,
 RA Waring A.J., To T., Fluharty C.B., Faull K.F.;
 RT "Preparation of the cerebroside sulfate activator (CSAct or saposin B)
 RT from human urine."
 RL Mol. Genet. Metab. 68:391-403(1999).
 RN [15]
 RP STRUCTURE OF CARBOHYDRATE ON ASN-215.
 RX MEDLINE=21110404; PubMed=11180632;
 RA Faull K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,
 RA Stevens R.L., Fluharty C.B., Fluharty A.L.;
 RT "Structure of the asparagine-linked sugar chains of porcine kidney and
 RT human urine cerebroside sulfate activator protein."
 RL J. Mass Spectrom. 35:1416-1424(2000).
 RN [16]
 RP SAPOSIN D DISULFIDE BONDS.
 RX MEDLINE=99337688; PubMed=10406958;
 RA Tatti M., Salvioli R., Ciaffoni F., Pucci P., Andolfo A.,
 RA Amoresano A., Vaccaro A.M.;
 RT "Structural and membrane-binding properties of saposin D."
 RL Eur. J. Biochem. 263:486-494(1999).
 RN [17]
 RP SAPOSIN B DISULFIDE BONDS.
 RX MEDLINE=22398398; PubMed=12510003;
 RA Ahn V.E., Faull K.F., Whitelegge J.P., Higginson J., Fluharty A.L.,
 RA Prive G.G.;
 RT "Expression, purification, crystallization, and preliminary X-ray
 RT analysis of recombinant human saposin B."
 RL Protein Expr. Purif. 27:186-193(2003).
 RN [18]
 RP MASS SPECTROMETRY.
 RC TISSUE=Urine;
 RX MEDLINE=99441404; PubMed=10510427;
 RA Faull K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,
 RA Krutchinsky A.N., Standing K.G., Waring A.J., Stevens R.L.,
 RA Fluharty C.B., Fluharty A.L.;
 RT "Cerebroside sulfate activator protein (Saposin B): chromatographic
 RT and electrospray mass spectrometric properties."
 RL J. Mass Spectrom. 34:1040-1054(1999).
 RN [19]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 195-273, AND MUTAGENESIS OF
 RP ILE-240.
 RX MEDLINE=22406333; PubMed=12518053;
 RA Ahn V.E., Faull K.F., Whitelegge J.P., Fluharty A.L., Prive G.G.;
 RT "Crystal structure of saposin B reveals a dimeric shell for lipid
 RT binding."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:38-43(2003).
 RN [20]
 RP REVIEW ON MLD VARIANTS.
 RX MEDLINE=95170731; PubMed=7866401;
 RA Gieselmann V., Zlotogora J., Harris A., Wenger D.A., Morris C.P.;
 RT "Molecular genetics of metachromatic leukodystrophy."
 RL Hum. Mutat. 4:233-242(1994).
 RN [21]
 RP VARIANT MLD ILE-217.
 RX MEDLINE=90147748; PubMed=2302219;
 RA Rafi M.A., Zhang X.-L., Degala G., Wenger D.A.;
 RT "Detection of a point mutation in sphingolipid activator protein-1
 RT mRNA in patients with a variant form of metachromatic
 RT leukodystrophy."
 RL Biochem. Biophys. Res. Commun. 166:1017-1023(1990).
 RN [22]
 RP SEQUENCE FROM N.A., AND VARIANT MLD ILE-217.
 RX MEDLINE=90207231; PubMed=2320574;

```
Query Match 100.0%; Score 2789; DB 1; Length 524;
Best Local Similarity 100.0%; Pred. No. 2e-176;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALFLASLLGALAGVLGKCTRGSAVWQVNTASDCAVKGKCLQTVWVKNKPTVKS 60
Db 1 MYALFLASLLGALAGVLGKCTRGSAVWQVNTASDCAVKGKCLQTVWVKNKPTVKS 60
Qy 61 LPCDICDVTAAAGDMLKDNATBEILLVLEKTCOWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICDVTAAAGDMLKDNATBEILLVLEKTCOWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPANIPILLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPANIPILLY 180
Qy 181 PQDGRSKPQKNGDVQDCIQMTDITQAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Db 181 PQDGRSKPQKNGDVQDCIQMTDITQAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Qy 241 CKNIYSQVSEIAIQMHMHQPKKEICALVGFCDVEKEMPOTLPAKVASKNVIPEALELVE 300
Db 241 CKNIYSQVSEIAIQMHMHQPKKEICALVGFCDVEKEMPOTLPAKVASKNVIPEALELVE 300
Qy 301 PIKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEECQEV 360
Db 301 PIKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEECQEV 360
Qy 361 VDTGSSILSLLEVSPELVCSMLHLCSTGRTRLPALTVHTVQKGGFCEVCKKLGVYLD 420
Db 361 VDTGSSILSLLEVSPELVCSMLHLCSTGRTRLPALTVHTVQKGGFCEVCKKLGVYLD 420
Qy 421 RNLEKNTKQBIILAELKGCSTFLPDPYQKQCDQFVAEVEPVLRIILVEWMDPSFVCLKIG 480
Db 421 RNLEKNTKQBIILAELKGCSTFLPDPYQKQCDQFVAEVEPVLRIILVEWMDPSFVCLKIG 480
Qy 481 ACPSAHPLLGTEKICWGPSYWCQNTETAACNAVEHCKRHVWN 524
Db 481 ACPSAHPLLGTEKICWGPSYWCQNTETAACNAVEHCKRHVWN 524

RESULT 2
SAP_BOVIN STANDARD; PRT; 525 AA.
ID -SAP BOVIN
AC P26779; Q9N2G4;
DT 01-AUG-1992 (Rel. 23, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Proactivator polypeptide precursor [Contains: Saposin A (Protein A);
DE Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside
DE sulfate activator) (CSAct) (Dispersin) (Sulfatide/GMI activator);
DE Saposin C (Co-beta-glucosidase) (Al activator) (Glucosylceramidase
DE activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D
DE (Protein C) (Component C)].
GN PSAP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OK NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS.
RC TISSU=Mammary gland;
RA Azuma N., Yoshida K.;
RT "RT-PCR cloning of bovine prosaposin.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 312-391.
RC TISSU=Spleen;
RA Sano A., Mizuno T., Kondoh K., Hineno T., Ueno S.-I., Kakimoto Y.,
RA Morita N.;
```

```
RT "Saposin-C from bovine spleen; complete amino acid sequence and
RT relation between the structure and its biological activity.";
RL Biochim. Biophys. Acta 1120:75-80(1992).
CC -!- FUNCTION: The lysosomal degradation of sphingolipids takes place
CC by the sequential action of specific hydrolases. Some of these
CC enzymes require specific low-molecular mass, non-enzymic proteins:
CC the sphingolipids activator proteins (coproteins) (By similarity).
CC -!- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC galactosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
CC Saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate.
CC -!- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GMI
CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
CC Saposin B forms a solubilizing complex with the substrates of the
CC sphingolipid hydrolases (By similarity).
CC -!- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
CC activator (EC 3.1.4.12) (By similarity).
CC -!- SUBUNIT: Saposin B is a homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- PTM: This precursor is proteolytically processed to 4 small
CC peptides, which are similar to each other and are sphingolipid
CC hydrolase activator proteins (By similarity).
CC -!- SIMILARITY: Contains 2 saposin A-type domains.
CC -!- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AB036791; BAA95677.1; -
DR InterPro; IPR003119; SapA.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SapB_1; 4.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 4.
KW Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat.
FT SIGNAL 1 16
FT PROPEP 17 58
FT CHAIN 60 142 SAPOSIN A.
FT PROPEP 144 195
FT CHAIN 196 275 SAPOSIN B.
FT PROPEP 277 310
FT CHAIN 312 392 SAPOSIN C.
FT PROPEP 393 404
FT CHAIN 406 487 SAPOSIN D.
FT PROPEP 489 525
FT DOMAIN 21 54
FT DOMAIN 59 142 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 194 276 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 312 393 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 406 487 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 492 525 SAPOSIN-LIKE TYPE B 4.
FT DISULFID 63 138 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 199 272 BY SIMILARITY.
FT DISULFID 202 266 BY SIMILARITY.
```

```

FT DISULFID 231 242 BY SIMILARITY.
FT DISULFID 316 389 BY SIMILARITY.
FT DISULFID 319 383 BY SIMILARITY.
FT DISULFID 347 358 BY SIMILARITY.
FT DISULFID 410 483 BY SIMILARITY.
FT DISULFID 413 477 BY SIMILARITY.
FT DISULFID 441 452 BY SIMILARITY.
FT CARBOHYD 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 427 427 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 127 127 H -> R.
FT VARIANT 260 263 MQPK -> IRIR.
FT VARIANT 317 317 E -> Q (IN REF. 2).
FT CONFLICT 317 317 R -> S (IN REF. 1).
FT CONFLICT 367 367
SQ SEQUENCE 525 AA; 58120 MW; 293AFC0FB9C4FA99 CRC64;

Query Match 87.0%; Score 2425.5; DB 1; Length 525;
Best Local Similarity 84.2%; Pred. No. 1.7e-152;
Matches 442; Conservative 45; Mismatches 37; Indels 1; Gaps 1;

Qy 1 MYAFLLASLLGALAGPVLGLKECTRGSAVWQVNVKTSADCGAVKHGICLTQVWNPVTYS 60
Db 1 MYSFVLSLGLGALASPLVIGRECTRGSAVWQVNVKTAADCGAVQHCLQTVWNSKPTVKS 60

Qy 61 LPDCICKOVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
Db 61 LPDCICKOVITAAGNLLKDNATEQEILMYLERTCDWLPKPNMSASCKEIVDSYLPVLDM 120

Qy 121 IKGMSRPGVCSALNLCESLOKHLAELNHQKLESNKIPELDMTEVVAPPFMANIPLLLY 180
Db 121 IKGMSHPGEVCSALNLCESLOKHLAELNHQKLESNQIPELDMAEVVAPPFMANIPELLY 180

Qy 181 PDGPRSRKPPKD-NGDVQCDCIQMVTDIQTAVRTNSTFVOALVEHVKEECDRLGPGMAD 239
Db 181 PDGSHSRKPPKANGVNCQDCIQLVTDVQELRTNSTFVEALVDHAKCEDRIGPGMSD 240

Qy 240 ICKNVISQYSEITAIOMHMQPKETCALVGFCDVEKPMQTLVPAKVASKNIVPALELV 299
Db 241 MCKNVINQYSEVAIQVMHMQPKETCVLAGFCDEKPMQTLVPAEAVSENVIPALGLV 300

Qy 300 EPIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILIDAPDKMSKLPKSLSECOE 359
Db 301 EPIKDPAPAKADIYCEVCEFLVKEVTKLIDNNKTEKEILIDAPDKMSKLPKSLSECOE 360

Qy 360 VVDYTGSSILILEEVSPELVCSMLHLCSGTRLPALT VHVTOPKGGFCFCEVCKLVGYL 419
Db 361 VVDYTGSSILILEEVSPELVCSMLHLCSGTRLPALT VHVTOPKGGFCFCEVCKLVGYL 420

Qy 420 DRNLEKSTKQEILAALEKGSFLPDVPYKOCQOFVAEYEPVLIETILVEWDPFSFVCLKI 479
Db 421 DRNLEKSTKQEILAALEKGSFLPDYRQKQCOQFVTEYEPVLIETILVEWDPFSFVCLKI 480

Qy 480 GACSAHKPLGTBKCITWSPSYWCNTTETAQCNNAVEHCKRHHVN 524
Db 481 GACPAHKPLLGAEKCVWSPSYWCNNESAAALCNNAVEHCHRRHVN 525

RESULT 3
SAP_RAT
ID SAP_RAT STANDARD; PRT; 554 AA.
AC P10560; 062841; 064190;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN PSAP OR SGP1
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

```

```

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Sertoli cells;
RX MEDLINE=89000647; PubMed=3048385;
RA Collard M.W., Sylvester S.R., Tauruta J.K., Griswold M.D.;
RT "Biosynthesis and molecular cloning of sulfated glycoprotein 1
RT secreted by rat Sertoli cells: sequence similarity with the
RL 70-kilodalton precursor to sulfatide/GM1 activator.";
RL Biochemistry 27:4557-4564 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96128541; PubMed=8573994;
RA Morales C.R., El-Alfy M., Zhao Q., Igdoura S.A.;
RT "Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in
RT Sertoli cells.";
RL Histol. Histopathol. 10:1023-1034 (1995).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RX MEDLINE=96175245; PubMed=8601692;
RA Morales C.R., El-Alfy M., Zhao Q., Igdoura S.A.;
RT "Expression and tissue distribution of rat sulfated glycoprotein-1
RT (prosaposin).";
RL J. Histochem. Cytochem. 44:327-337 (1996).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: Contains 2 saposin A-type domains.
CC -1- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; M19936; AAA42136.1; -
CC EMBL; S81353; AAB36042.2; -
CC EMBL; S81373; AAB36233.2; -
CC PIR; A28716; A28716.
CC InterPro; IPR003119; Sapa.
CC InterPro; IPR007856; SapaB.
CC InterPro; IPR008138; SapaB.2.
CC InterPro; IPR008140; SapaB.sub.
CC InterPro; IPR008373; Saposin.
CC InterPro; IPR008139; SaposinB.
CC Pfam; PF02199; SAPA.2.
CC Pfam; PF05184; SapaB.1; 3.
CC Pfam; PF03489; SapaB.2; 4.
CC PRINTS; PR01797; SAPOSIN.
CC ProDom; PD001732; SapaB.sub.3.
CC SMART; SM00162; SAPA.2.
CC SMART; SM00118; SAPA.4.
KW Signal; Glycoprotein; Repeat.
FT SIGNAL 1 16
FT CHAIN 17 554 SULFATED GLYCOPROTEIN 1.
FT DOMAIN 21 54 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 61 138 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 193 274 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 310 391 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 435 516 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 521 554 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 197 270 BY SIMILARITY.
FT DISULFID 200 264 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 314 387 BY SIMILARITY.
FT DISULFID 317 381 BY SIMILARITY.
FT DISULFID 345 356 BY SIMILARITY.
FT DISULFID 439 512 BY SIMILARITY.
FT DISULFID 442 506 BY SIMILARITY.

```



```
FT DISULFID 470 481 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 115 115 P -> L (IN REF. 2 AND 3).
FT CONFLICT 299 299 D -> E (IN REF. 2).
FT CONFLICT 462 462 I -> V (IN REF. 3).
FT CONFLICT 527 527 W -> R (IN REF. 3).
FT CONFLICT 536 536 S -> M (IN REF. 3).
SQ SEQUENCE 554 AA; 61123 MW; DFE3F3A3A0520C6B CRC64;

Query Match 71.6%; Score 1996; DB 1; Length 554;
Best Local Similarity 66.3%; Pred. No. 3,56-124;
Matches 368; Conservative 77; Mismatches 78; Indels 32; Gaps 3;

Qy 1 MYALFLIASLLGALAGLVGLKTRGSAVWQNVKTSADCGAVKHCLQTVNKNPTVKS 60
Db 1 MYALALLASLLVLTALTSVPQDPKICSGSVAVVCNDVKTAVDCRAVRHCQQMWKSKPTAKS 60

Qy 61 LPCDICKDVVTAAGDMLKDNATEBEILVYLEKTCMDLKPKNMGSACKEIIVDSYLPVLIDI 120
Db 61 LPCDICKTVVTEAGNLLKDNATEBEILVYLEKTCAMIHDSLSLSCKEVDSYLPVLIDM 120

Qy 121 IKGEMSPGVCNAINCESLQKHLAEINHQLESNKIPELDMTEVVPFMANIPILLY 180
Db 121 IKGEMSPGVCNAINCQSLQYLAQN-QRQLESNKIPEVDLARVVAFPMSPNIPILLY 179

Qy 181 PDGPRSKPQKNGVDCQICQIMVTDIOTAVRTNSTFVQALVHVHKEECDRLPGNADI 240
Db 181 PDGPRSKPQKNGVDCQICQIMVTDIOTAVRTNSTFVQALVHVHKEECDRLPGVSDI 239

Qy 241 CKNYISQYSIAIOMMHMOPKEICALVGFCDVEKEMPMOTLVPAKVASKNVIPALSELVE 300
Db 240 CKNYVDQYSEVAVQMMHMQPKEICVMVGFCDVEKVRPMETLVPAEAIKNIPALELTD 299

Qy 301 PIKHEVPAKSDVYCEVFLVKEVTKLIDNNKTEKEILDAFDKMSKLPKSISEECQEV 360
Db 300 PYEQDVIAQNVIPQVQCVLMRKLSELLINNAATELLIKGLSKACSLLPAPASTKQEV 359

Qy 361 VDTYGSILSLLEEVSPVCSMLHLCSG-----TRLP----- 394
Db 360 LVTFGSLLDVLMHEVNPFLCGVISCSANPNLVGTLEQPAAVSALPKPEAPPKQPE 419

Qy 395 -----ALTVHTVTPKDCGCEVCKLVGYLDRLNLEKNSTKQEIILAALEKGCSPDPDYQK 449
Db 420 EPKQALRAHVPPQNGGFCVCKLVLYEHLNLEKNSTKEEILAALEKGCSPDPDYQK 479

Qy 450 QCOFVAEYFVLIEILVEVMDPSFVCLKIGACPSAHKPLLGTEKCIWGPSYMCQNTETA 509
Db 480 QCOFVAEYFVLIEILVEVMDPSFVCSKIGVCPKSAVKLLLGTEKCVWGPYMCQNTETA 539

Qy 510 AOCNAVEHCKRHVWN 524
Db 540 ARCNVADHCKRHVWN 554

RESULT 4
SAP_MOUSE
ID SAP_MOUSE STANDARD; PRT; 557 AA.
AC Q61207; Q60861; Q64006; Q64219;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN PSAP OR SGP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
```

```
RX MEDLINE=92272718; PubMed=1590788;
RA Tada M., Sakiyama T., Endo H., Kitagawa T.;
RT "The primary structure of mouse saposin.";
RL Biochem. Biophys. Res. Commun. 184:1266-1272(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94272317; PubMed=8003952;
RA Sprecher-Levy H., Orr-Urtreger A., Ioni P., Horowitz M.;
RT "Murine prosaposin: expression in the reproductive system of a gene
implicated in human genetic disease.";
RL Cell. Mol. Biol. 40:233-233(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96084310; PubMed=8565332;
RA Cao Q.P., Crain W.R.;
RT "Expression of SGP-1 mRNA in preimplantation mouse embryos.";
RL Dev. Genet. 17:263-271(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Zhao Q.Q., Hay N.N., Morales C.R.;
RT Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: Contains 2 saposin A-type domains.
CC -1- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; S36200; AAB22175.1; -
DR EMBL; S71616; AAB31059.1; -
DR EMBL; U27340; AAA92567.1; -
DR EMBL; U57999; AAB02695.1; -
DR PIR; JH0604; JH0604.
DR MGD; MGI:97783; Psap.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SapB_1; 4.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 4.
DR Signal; Glycoprotein; Repeat.
KW SIGNAL
FT CHAIN 1 16 BY SIMILARITY.
FT SIGNAL 1 16 SULFATED GLYCOPROTEIN 1.
FT CHAIN 17 57 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 21 54 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 59 142 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 193 277 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 313 394 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 438 519 SAPOSIN-LIKE TYPE A 2.
FT DOMAIN 524 557 BY SIMILARITY.
FT DISULFID 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 197 273 BY SIMILARITY.
FT DISULFID 200 267 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 317 390 BY SIMILARITY.
FT DISULFID 320 384 BY SIMILARITY.
FT DISULFID 348 359 BY SIMILARITY.
FT DISULFID 442 515 BY SIMILARITY.
FT DISULFID 445 509 BY SIMILARITY.
```

Query Match
Best Local Similarity 69.5%; Score 1938.5; DB 1; Length 557;
Matches 355; Conservative 78; Mismatches 90; Indels 35; Gaps 3;

473 484 BY SIMILARITY.
80 N-LINKED (GLCNAC. . .) (POTENTIAL).
214 N-LINKED (GLCNAC. . .) (POTENTIAL).
334 N-LINKED (GLCNAC. . .) (POTENTIAL).
459 N-LINKED (GLCNAC. . .) (POTENTIAL).
83 Q -> E (IN REF. 2).
158 I -> V (IN REF. 3).
160 MISSING (IN REF. 2).
171 MS -> SA (IN REF. 3).
172 V -> L (IN REF. 2).
244 M -> I (IN REF. 3).
254 L -> W (IN REF. 2).
255 MISSING (IN REF. 3).
262 N -> D (IN REF. 2).
307 F -> L (IN REF. 2).
322 AL -> GV (IN REF. 1).
349 G -> D (IN REF. 3).
367 L -> Q (IN REF. 2).
370 I -> T (IN REF. 3).
373 A -> T (IN REF. 3).
391 R -> L (IN REF. 3).
393 A -> R (IN REF. 2 AND 3).
406 P -> R (IN REF. 2).
430 C -> F (IN REF. 3).
445 L -> P (IN REF. 4).
448
557 AA; 61422 MW; 134593E20499E35E CRC64;

Query Match
Best Local Similarity 69.5%; Score 1938.5; DB 1; Length 557;
Matches 355; Conservative 78; Mismatches 90; Indels 35; Gaps 3;

473 484 BY SIMILARITY.
80 N-LINKED (GLCNAC. . .) (POTENTIAL).
214 N-LINKED (GLCNAC. . .) (POTENTIAL).
334 N-LINKED (GLCNAC. . .) (POTENTIAL).
459 N-LINKED (GLCNAC. . .) (POTENTIAL).
83 Q -> E (IN REF. 2).
158 I -> V (IN REF. 3).
160 MISSING (IN REF. 2).
171 MS -> SA (IN REF. 3).
172 V -> L (IN REF. 2).
244 M -> I (IN REF. 3).
254 L -> W (IN REF. 2).
255 MISSING (IN REF. 3).
262 N -> D (IN REF. 2).
307 F -> L (IN REF. 2).
322 AL -> GV (IN REF. 1).
349 G -> D (IN REF. 3).
367 L -> Q (IN REF. 2).
370 I -> T (IN REF. 3).
373 A -> T (IN REF. 3).
391 R -> L (IN REF. 3).
393 A -> R (IN REF. 2 AND 3).
406 P -> R (IN REF. 2).
430 C -> F (IN REF. 3).
445 L -> P (IN REF. 4).
448
557 AA; 61422 MW; 134593E20499E35E CRC64;

Query Match
Best Local Similarity 69.5%; Score 1938.5; DB 1; Length 557;
Matches 355; Conservative 78; Mismatches 90; Indels 35; Gaps 3;

473 484 BY SIMILARITY.
80 N-LINKED (GLCNAC. . .) (POTENTIAL).
214 N-LINKED (GLCNAC. . .) (POTENTIAL).
334 N-LINKED (GLCNAC. . .) (POTENTIAL).
459 N-LINKED (GLCNAC. . .) (POTENTIAL).
83 Q -> E (IN REF. 2).
158 I -> V (IN REF. 3).
160 MISSING (IN REF. 2).
171 MS -> SA (IN REF. 3).
172 V -> L (IN REF. 2).
244 M -> I (IN REF. 3).
254 L -> W (IN REF. 2).
255 MISSING (IN REF. 3).
262 N -> D (IN REF. 2).
307 F -> L (IN REF. 2).
322 AL -> GV (IN REF. 1).
349 G -> D (IN REF. 3).
367 L -> Q (IN REF. 2).
370 I -> T (IN REF. 3).
373 A -> T (IN REF. 3).
391 R -> L (IN REF. 3).
393 A -> R (IN REF. 2 AND 3).
406 P -> R (IN REF. 2).
430 C -> F (IN REF. 3).
445 L -> P (IN REF. 4).
448
557 AA; 61422 MW; 134593E20499E35E CRC64;

KW Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat;
KW GM2-gangliosidosis.

	SIGNAL	1	17	POTENTIAL.
FT	PROPEP	18	60	
FT	CHAIN	61	143	SAPOSIN A.
FT	PROPEP	145	193	
FT	CHAIN	194	276	SAPOSIN B.
FT	PROPEP	278	305	
FT	CHAIN	307	387	SAPOSIN C.
FT	PROPEP	389	398	
FT	CHAIN	399	480	SAPOSIN D.
FT	PROPEP	482	518	
FT	DOMAIN	22	55	SAPOSIN-LIKE TYPE A 1.
FT	DOMAIN	60	143	SAPOSIN-LIKE TYPE B 1.
FT	DOMAIN	193	277	SAPOSIN-LIKE TYPE B 2.
FT	DOMAIN	307	388	SAPOSIN-LIKE TYPE B 3.
FT	DOMAIN	399	480	SAPOSIN-LIKE TYPE B 4.
FT	DOMAIN	485	518	SAPOSIN-LIKE TYPE A 2.
FT	DISULFID	64	139	BY SIMILARITY.
FT	DISULFID	67	133	BY SIMILARITY.
FT	DISULFID	95	107	BY SIMILARITY.
FT	DISULFID	197	273	BY SIMILARITY.
FT	DISULFID	200	267	BY SIMILARITY.
FT	DISULFID	229	240	BY SIMILARITY.
FT	DISULFID	311	384	BY SIMILARITY.
FT	DISULFID	314	378	BY SIMILARITY.
FT	DISULFID	342	353	BY SIMILARITY.
FT	DISULFID	403	476	BY SIMILARITY.
FT	DISULFID	406	470	BY SIMILARITY.
FT	DISULFID	434	445	BY SIMILARITY.
FT	CARBONYD	81	81	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBONYD	214	214	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBONYD	328	328	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBONYD	420	420	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CONFLICT	94	94	R -> T (IN REF. 2).
FT	CONFLICT	486	486	E -> D (IN REF. 2).
FT	SEQUENCE	518 AA;	57601 MW;	B803000E891C3963 CRC64;

Query Match 60.2%; Score 1679.5; DB 1; Length 518;
 Best Local Similarity 57.4%; Pred. No. 2.2e-103;
 Matches 301; Conservative 96; Mismatches 114; Indels 13; Gaps 5;

Qy	4	LFLIASLIGALAGPVLGKECTRGSAVWCVNQVKTASDCGAIVKCLQTWNKPVTKSIPC	63
Dd	5	LTLTLLGAAAASAPFVLWQKDCAKGPEVWCOSLTASQCQGAIVKHQQNWSKPAVNSIPC	64
Qy	64	DICKDVTAAGDMLKDNATEBEEILVLEKTCMDPLPKPNMSASCKEIVDSPVLVIDIKG	123
Dd	65	DLCELVTVGVKVLKNGTEDEIRSYLEKRCFEPLPDQGLASECKEIVDSYLPVIMDKIE	124
Qy	124	EMSRPGVCSSALNICESLQKHLAEINHQQLSESNKIPELDMTEVVVAFFMANIPILLYPOD	183
Dd	125	EFDKPEVVCSSALSQSLOKHLAAAMKLQQLQSNKIPELDFSELTPPMANVPILLYPOD	184
Qy	184	GPRSKPQKONGDYVCODCIOMVTDIOATVRNTSFQVALVEHVKEECDRLGPGMADICKN	243
Dd	185	KPKQ--SKATEDVCQDCIRLUITVDQEAVRTNFVKSLVAHAKEECDRLGPGSDMCKS	242
Qy	244	YISOYSRIAIQMNMHM---QPKEICALGVGDEVKEMPMTQTLVPAKVASKNVIPALELVE	300
Dd	243	YISEYSDLAIQNMHMKDDQKDICAMVGFCPSVKSVPLQTLVPAQVHE-----VQME	296
Qy	301	PIKGEHPAKSDVYCEVCFELVKEVTKLIIDNNKTEKTELDAFPDKMCSKLPKLSLSECOEV	360
Dd	297	TVEKATVQEKTFSCVEICETVMKEVTVGLSENKTEETEEIVHEMVVVCYLLPASVKDQCKDF	356
Qy	361	VDTYGSSILSLLEVSPELVCSMLHLCSGTRLPALTVHTVTPKDGGFCEVCKKLVGYLD	420
Dd	357	IENVGOALIDMLELATNPFAVCVWLKKCANPPQQPV-VVKPA-GGFCDICKMLVAYD	414
Qy	421	RNLKNSTQBFIILAALKGCSSFDPYPQOKDCQDFVAIEYEPVLIELIVEMDPSPFVCLKIG	480
Dd	415	KELSKNATTTBIALLEKVCHFLPESVSDOCVFVEQVEPVVVOLLAEMBPPTFYCTKLG	474

481 APCSAHKPLLTGTEKICWSPVWQNTETAAQCAVHCKRHVWN 524
| : | | | | : | : | | | : | : | | | | : | : | | | |
475 VCGAAKPLLGEDACVWGPGYWCXKNMETAAQCAVHCKRHVWN 518

RESULT 6

SAP_PIG	STANDARD;	PRT;	80 AA.
ID -SAP_PIG			
AC	P81405;		
DT	15-DEC-1998 (Rel. 37, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Saposin B (Cerebroside sulfate activator) (CS-ACT) (Non-specific		
DE	activator) (Sphingolipid activator protein 1) (SAP-1).		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
OX	NCBI_TaxID=9823;		
RP	[1]		
RP	SEQUENCE OF 1-79.		
RC	TISSUE=Kidney;		
RC	MEDLINE=93229506; PubMed=8471613;		
RA	Stevens R.L., Faull K.F., Conklin K.A., Green B.N., Fluharty A.L.;		
RA	"Porcine cerebroside sulfate activator: further structural		
RT	characterization and disulfide identification.";		
RL	Biochemistry 32:4051-4059(1993).		
RN	[2]		
RP	SEQUENCE OF 1-64.		
RP	TISSUE=Kidney;		
RC	MEDLINE=92222651; PubMed=1562358;		
RA	Fluharty A.L., Katona Z., Meek W.B., Frei K., Fowler A.V.;		
RA	"The cerebroside sulfate activator from pig kidney: purification and		
RT	molecular structure.";		
RL	Biochem. Med. Metab. Biol. 47:66-85(1992).		
RN	[3]		
RP	STRUCTURE OF CARBOHYDRATE ON ASN-21.		
RC	MEDLINE=21110404; PubMed=11180632;		
RA	Faull K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,		
RA	Stevens R.L., Fluharty C.B., Fluharty A.L.;		
RT	"Structure of the asparagine-linked sugar chains of porcine kidney and		
RT	human urine cerebroside sulfate activator protein.";		
RL	J. Mass Spectrom. 35:1416-1424(2000).		
RN	[4]		
RP	MASS SPECTROMETRY.		
RC	TISSUE=Kidney;		
RC	MEDLINE=99441404; PubMed=10510427;		
RA	Faull K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,		
RA	Krutchinsky A.N., Standing K.G., Waring A.J., Stevens R.L.,		
RA	Fluharty C.B., Fluharty A.L.;		
RT	"Cerebroside sulfate activator protein (Saposin B): chromatographic		
RT	and electrospray mass spectrometric properties.";		
RL	J. Mass Spectrom. 34:1040-1054(1999).		
CC	-1- FUNCTION: Saposin B stimulates the hydrolysis of galacto-		
CC	cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1		
CC	gangliosides by beta-galactosidase (EC 3.2.1.23) and		
CC	globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).		
CC	Saposin B forms a solubilizing complex with the substrates of the		
CC	sphingolipid hydrolases.		
CC	-1- SUBUNIT: Saposin B is a homodimer (By similarity).		
CC	-1- PTM: The one residue extended Saposin B-Val is only found in a		
CC	minority of the chains.		
CC	-1- SIMILARITY: Contains 1 saposin B-type domain.		
DR	GlycoSuiteDB; P81405; -.		
DR	InterPro; IPR007856; SapB_1.		
DR	InterPro; IPR008138; SapB_2.		
DR	InterPro; IPR008373; Saposin.		
DR	InterPro; IPR008139; SaposinB.		
DR	Pfam; PF05184; SapB_1; 1.		
DR	Pfam; PF03489; SapB_2; 1.		
DR	PRINTS; PR01797; SAPOSIN.		
DR	SMART; SM00118; SAPB; 1.		
KW	Glycoprotein; Sphingolipid metabolism.		

```

FT CHAIN 1 79 SAPOSIN B.
FT CHAIN 1 80 SAPOSIN B-VAL.
FT DOMAIN 1 80 SAPOSIN-LIKE TYPE B.
FT DISULFID 4 77
FT DISULFID 7 71
FT DISULFID 36 47
FT CARBOHYD 21 21
SQ SEQUENCE 80 AA; 8949 MW; EFTBA249B63E789C CRC64;

Query Match 14.1%; Score 393; DB 1; Length 80;
Best Local Similarity 88.8%; Pred. No. 1.1e-19;
Matches 71; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 195 GDVQCDCIQMVTDTIQTAVRTNSTFVQALVHVKECRDLRGLPGMADICKNYSQYSEIAIQ 254
Db 1 GDVQCDCIQMVTDTIQTAVRTNSTFVQALVHVKECRDLRGLPGMADICKNYSQYSEIAIQ 60

QY 255 MMHMQPKFICGLVGFCEV 274
Db 61 MMHMQPKFICGLVGFCEV 80

RESULT 7
PSPB_MOUSE
ID PSPB_MOUSE STANDARD; PRT; 377 AA.
AC P50405;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pulmonary surfactant-associated protein B precursor (SP-B) (Pulmonary
DE surfactant-associated proteolipid SPL(Phe)).
DE SFTPB OR SFTPB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2J; TISSUE=Liver;
RX MEDLINE=95208782; PubMed=7900819;
RA Bruno M.A., Bohinski R.J., Carter J.E., Foss K.A., Whitsett J.A.;
RT "Structure and function of the mouse surfactant protein B gene.";
RL Am. J. Physiol. 268:1381-1389(1995).
CC -!- FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE
CC ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-
CC LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES
CC THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS
CC PER METER.
CC -!- SUBUNIT: Homodimer; disulfide-linked.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
CC protein. There are 4 surfactant-associated proteins: 2 collagenous,
CC carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small
CC hydrophobic proteins (SP-B and SP-C).
CC -!- SIMILARITY: Contains 1 saposin A-type domain.
CC -!- SIMILARITY: Contains 3 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S78114; AAB34846.2; -
DR HSP; P07988; IDFW.
DR MGD; MGI:109516; Sftpb.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapB.1.
DR InterPro; IPR008138; SapB.2.
DR InterPro; IPR008140; SapB_sub.

```

```

DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR InterPro; IPR008137; Surfactant_B.
DR Pfam; PF02199; SAPA; 1.
DR Pfam; PF05184; SapB.1; 1.
DR Pfam; PF03489; SapB.2; 3.
DR PRINTS; PR01797; SAPOSIN.
DR PRODOM; PD001732; SapB_sub; 1.
DR PRODOM; PD008002; Surfactant_B; 1.
DR SMART; SM00162; SAPA; 1.
DR SMART; SM00118; SAPP; 3.
KW Surface film; Gaseous exchange; Glycoprotein; Repeat.
FT PROPEP 1 191 PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT CHAIN 192 270 B.
FT PROPEP 271 377
FT DOMAIN 27 60 SAPOSIN-LIKE TYPE A.
FT DISULFID 199 268 BY SIMILARITY.
FT DISULFID 202 262 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 239 239 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 307 307 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 377 AA; 41728 MW; CB687A82BA3FC56C CRC64;

Query Match 13.7%; Score 382; DB 1; Length 377;
Best Local Similarity 22.9%; Pred. No. 4e-18;
Matches 118; Conservative 75; Mismatches 154; Indels 168; Gaps 18;

QY 4 LFLASLL--GAALAGPVLGLKECTRGSAVWQVKTASDCGAVKHCIQTVMNKPTVKSL 61
Db 10 LLLFLTLCPGAAITS--ASSLECAQPGQFQCQSLHVAQCRAHGCLQEVMGHAGANDL 67

QY 62 PCDCIKDVVTAAGMDLKDNAMEEILVYLEKTCWLPKPNMSASCKEIVDSVLPVLDII 121
Db 68 -QCECEDIVHLLTKTKEDAFQEARKELEOECDDLPLKLVPRCQVDVLPVLDIYF 126

QY 122 KGEMSRPGVCNALNCESLOKHLAELNHQKLESNKIPELDMTEVAVFMANIPLLLYP 181
Db 127 QSQIN-PKAINHVGLC----- 142

QY 182 QDGRSKPQKNDGVDVQCDCIQMVTDTIQTAVRTNSTFVQALVHVKECRDLRGLPGMADIC 241
Db 143 ---PRGAQKPEQN-----PGMPD-- 157

QY 242 KNYISQYSEIAIQMMHMQPKFICGLVGFCEVDEKEMPMQTLVPAKVASKNVIPALE---L 298
Db 158 -----AVPNPLDLKLVLPVLPALL 177

QY 299 VEP-----IKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKIELD-AFDKMCCKLPK 351
Db 178 ARPGHTQDFSEQLPIPLP-FCWLCTRLTKRQAVI-----PKGVLAVALVQVCHVVP 231

QY 352 SLSECEQVVDVYSSILSILLESPELVCSMLHLC-----GTRLPALTVMHTQ-PKD 405
Db 232 VVGICQCCLAEYTVLLDALLGRVVPQLVGLVLCSTEDAMGPALPAVEPLIEEPLQ 291

QY 406 GGFCEVCKLVGLDRNLKNSKTKQELAALEKGC-SFLPDYQKQCDQFAVEPVLE 464
Db 292 DTECHFCRSVI-----NOAWNTSEQAMPQAMQACLRFLWD--RQCEQFVEQHMPQLLA 344

QY 465 ILVEWMDPSFVCLKIGACPSAHKPLILGTEKCIWGP 499
Db 345 LVPRSDAHITCQALGVCEAPASPL-----QCFTP 375

RESULT 8
PSPB_RAT
ID PSPB_RAT STANDARD; PRT; 376 AA.
AC P22355;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pulmonary surfactant-associated protein B precursor (SP-B) (Pulmonary

```


RC TISSUE=Brain; PubMed=12477932;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg H., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Tohiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Richards S., Worley K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [5]
RP SEQUENCE OF 99-381 FROM N.A.
RX MEDLINE=87231940; PubMed=3035561;
RA Glasser S.W., Korfhagen T.R., Weaver T., Pilot-Matias T., Fox J.L.,
RA Whitsett J.A.;
RT "cDNA and deduced amino acid sequence of human pulmonary surfactant-
associated proteolipid sPL(Ph)".
RL Proc. Natl. Acad. Sci. U.S.A. 84:4007-4011 (1987).
RN [6]
RP SEQUENCE OF 139-381 FROM N.A.
RX MEDLINE=88139786; PubMed=3343343;
RA Revak S.D., Merritt T.A., Degryse E., Stefani L., Courtney M.,
RA Hallman M., Cochran C.G.;
RT "Use of human surfactant low molecular weight apoproteins in the
reconstitution of surfactant biologic activity.";
RL J. Clin. Invest. 81:826-833 (1988).
RN [7]
RP SEQUENCE OF 201-279, AND DISULFIDE BONDS.
RX MEDLINE=92233937; PubMed=1568474;
RA Johansson J., Joernvall H., Curstedt T.;
RT "Human surfactant polypeptide SP-B. Disulfide bridges, C-terminal
end, and peptide analysis of the airway form.";
RL FEBS Lett. 301:165-167 (1992).
RN [8]
RP STRUCTURE BY FTIR OF 201-225.
RX MEDLINE=20256237; PubMed=10798379;
RA Gordon L.M., Lee K.Y., Lipp M.M., Zasadzinski J.A., Walther F.J.,
RA Sherman M.A., Waring A.J.;
RT "Conformational mapping of the N-terminal segment of surfactant
protein B in lipid using 13C-enhanced Fourier transform infrared
spectroscopy.";
RL J. Pept. Res. 55:330-347 (2000).
RN [9]
RP VARIANT ILE-131.
RX MEDLINE=20525299; PubMed=11076040;
RA Lin Z., Pearson C., Chinchilli V., Pietschmann S.M., Luo J., Pison U.,
RA Floros J.;
RT "Polymorphisms of human SP-A, SP-B, and SP-D genes: association of
SP-B Thr311le with ARDS.";
RL Clin. Genet. 58:181-191 (2000).
CC -1- FUNCTION: Pulmonary surfactant-associated proteins promote
alveolar stability by lowering the surface tension at the air-
liquid interface in the peripheral air spaces. SP-B increases
the collapse pressure of palmitic acid to nearly 70 millineutons
per meter.
CC -1- SUBUNIT: Homodimer; disulfide-linked.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- POLYMORPHISM: There seems to be an association between the Ile-131
polymorphism and acute respiratory distress syndrome (ARDS).
CC -1- DISEASE: Defects in SFTPB are a cause of congenital pulmonary
alveolar proteinosis (PAP) [MIM:265120]. PAP is an autosomal

CC recessive fatal respiratory disease.
CC -1- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
protein. There are 4 surfactant-associated proteins: 2 collagenous,
CC carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small
CC hydrophobic proteins (SP-B and SP-C).
CC -1- SIMILARITY: Contains 1 saposin A-type domain.
CC -1- SIMILARITY: Contains 3 saposin B-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M24461; AAB59541.1; --
CC EMBL; J02761; AAA60212.1; --
CC EMBL; AF400074; AAK77913.1; --
CC EMBL; BC032785; AAH32785.1; --
CC EMBL; M16764; AAA88099.1; ALT_INIT.
CC EMBL; M19097; AAA36628.1; --
CC PIR; A31361; LNHUB.
CC PDB; 1DFW; 24-OCT-01.
CC PDB; 1KWE; 07-AUG-02.
CC Genew; HGNC:10801; SFTPB.
CC MIM; 178640; --
CC GO; GO:0007397; P:histogenesis and organogenesis; TAS.
CC GO; GO:0007585; P:respiratory gaseous exchange; TAS.
CC InterPro; IPR003119; Sapa.
CC InterPro; IPR007856; Sapa_1.
CC InterPro; IPR008138; Sapa_2.
CC InterPro; IPR008140; Sapa_sub.
CC InterPro; IPR008373; Saposin.
CC InterPro; IPR008139; SaposinB.
CC InterPro; IPR008137; Surfactant_B.
CC Pfam; PF02199; SAPA; 1.
CC Pfam; PF05184; Sapa_1; 1.
CC Pfam; PF03489; Sapa_2; 3.
CC PRINTS; PRO1797; SAPOSIN.
CC ProDom; PD001732; Sapa_sub; 1.
CC ProDom; PD008002; Surfactant_B; 1.
CC SMART; SM00162; SAPA; 1.
CC SMART; SM00118; Sapa; 3.
KW Surface film; Gaseous exchange; Glycoprotein; Repeat; Polymorphism;
KW 3D-structure.
FT PROPEP 1 200
FT CHAIN 201 279
FT
FT PROPEP 280 381
FT DOMAIN 28 61
FT DISULFID 208 277
FT DISULFID 211 271
FT DISULFID 235 246
FT DISULFID 248 248
FT CARBOHYD 129 129
FT CARBOHYD 311 311
FT VARIANT 131 131
FT VARIANT 176 176
FT VARIANT 228 228
FT VARIANT 228 228
FT VARIANT 272 272
FT CONFLICT 178 178
FT CONFLICT 318 318
FT HELIX 208 221
SQ SEQUENCE 381 AA; 42117 MW; 9FD7F66678A35153 CRC64;
B.
SAPOSIN-LIKE TYPE A.
INTERCHAIN.
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
T -> I.
/FTid=VAR_006948.
L -> F.
/FTid=VAR_013099.
A -> R.
/FTid=VAR_006949.
A -> I.
/FTid=VAR_006950.
R -> H.
/FTid=VAR_013100.
L -> V (IN REF. 6).
P -> L (IN REF. 5).
P

```
Query Match      12.5%; Score 349.5; DB 1; Length 381;
Best Local Similarity 21.6%; Pred. No. 5.6e-16;
Matches 111; Conservative 72; Mismatches 171; Indels 161; Gaps 16;

QY 2 YALFLLASLGAALAGPVGLGKCTRGSAVWQNVKTSADCGAVKHCLQTVWKNKPTVKS 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 WLLLLLPLCGPTAAWTTSSLAQGGPFWCQSLQALQCRALGHCLQEVWGHVGGADD 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 PCDICKDVVTAAGDMLKDNATEEILVYLEKTDWLPKPNMSASCKEIVDSYLPVILDI 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 -CQECEDIVHLNKWAKEAFQDTRKFELEQECNVLPKLLMPQCNOVLDDYFFLVIDYF 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 122 KGEISRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYP 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 QNQTDSNG-ICMHLGLCKSRQ----- 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 QDGRPSRPPKPDNGDVCQDCIQMTDITQAVRTNSTFTVQALVEHVKECDRLGPGMADIC 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 148 -----PEPEQE----- 158
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 KNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPQTLVPAKVASKNVIPALELVEP 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 159 -----PLPK----- 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 302 IKGHEVPAKSDV-----YCEVCEFLVKEVTKLIDNNKTEKEILDAPDKMCKSLPK 351
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 185 LQARPGPHTDLSRQGFPIPLPYCWLCLRALIKRIQAMIPKG-----ALAVAQVQCRVVP 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 352 SLSECEQVVDVTVSSILSLILEVSPVLCVSMHLCS-----CTRLPALTVHVTQPKDG 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 VAGIGICQCLAEYRVSIVLLDTLGLRMLPOLVCLRLVLRCSMDSDSAGPRSP---TGEWLP 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 407 GFCEVCKKLYGLDRNLKNSKTEKILAALEKGC--SFLPDYKQCDQFVAEYEPVLIE 464
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 E-CHLCSVNT-----TOAGNSSEQAIQAMLOACVGSWLD---REKCKQFVEQHTPOLL 348
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 465 ILVEVMDPSFVCLKIGACPSAKHPLGTEKICWGP 499
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 349 LVPRGWDHATTTCQALGCVGTMSPL-----QCIHSP 379
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
PSPB_CANFA STANDARD; PRT; 363 AA.
AC P17129;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pulmonary surfactant-associated protein B precursor (SP-B) (6 kDa
DE protein) (Pulmonary surfactant-associated protein B) (Phe)
DE (Pulmonary surfactant-associated protein 18) (SP 18) (Fragment).
GN SFTPB OR SFTPB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 182-211.
RC TISSUE=Lung;
RA MEDLINE=87092398; PubMed=3467361;
RA Hawgood S., Benson B.J., Schilling J., Damm D., Clements J.A.,
RA White R.T.;
RT "Nucleotide and amino acid sequences of pulmonary surfactant protein
RT SP 18 and evidence for cooperation between SP 18 and SP 28-36 in
RT surfactant lipid adsorption."
RL Proc. Natl. Acad. Sci. U.S.A. 84:66-70(1987).
CC -!- FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE
CC ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-
CC LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES
CC THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS
CC PER METER.
CC -!- SUBUNIT: Homodimer; disulfide-linked.
CC -!- SUBCELLULAR LOCATION: Extracellular.
```

Db 185 -----LP-----YCMCLRTLIKRIQAMI-----PKGVLAVTVQVCHVPLVVGICQC 228
Qy 360 VVDYGGSSILSILEEVSPELVCSMLHLS-----GTRLPALTVHTVTPKDGCFCEVCKK 414
Db 229 LGERYTVLLDALLGRMLPOLVGLVLRCSHDSAGPALASPSE--WSPQESK--CQLQWF 286
Qy 415 LVGYLDNRNLEKSTKOBILAALEKGC--SFLPPYQKQCDQFVAEVEPVLILVEMVDP 472
Db 287 VT-----TOAGNHSEQATPAIRQACLSWLD---RQKCEQFVEQHMPRLQTLASGRDA 338
Qy 473 SFVCLKIGACPSAHKPLGTKEKINGPSY 501
Db 339 HTTCQALGACRTTFSPL-----QCIHPIHF 363

RESULT 11
PSPB RABIT
ID PSPB RABIT STANDARD; PRT; 370 AA.
AC P15285;
DT 01-APR-1990 (Rel. 14, Created)
DT 16-OCT-2001 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pulmonary surfactant-associated protein B precursor (SP-B) (6 kDa
DE protein) (Pulmonary surfactant-associated proteolipid SPL(Phe)).
GN SFTPB OR SFTPB.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RN [1] SEQUENCE FROM N.A.
RN [1] TISSUE=lung;
RN [1] MEDLINE=89228033; PubMed=2469419;
RA Xu J., Richardson C., Ford C., Spencer T., Li-Juan Y., Mackie G.,
RA Hammond G., Possmayer F.;
RT "Isolation and characterization of the cDNA for pulmonary surfactant-
RT associated protein-B (SP-B) in the rabbit.";
RT Biochem. Biophys. Res. Commun. 160:325-332(1989).
RN [2]
RN [2] SEQUENCE FROM N.A.
RN [2] STRAIN=New Zealand white;
RN [2] MEDLINE=95208794; PubMed=7900830;
RA Margana R.K., Boggaram V.;
RT "Transcription and mRNA stability regulate developmental and hormonal
RT expression of rabbit surfactant protein B gene.";
RL Am J. Physiol. 268:L481-L490(1995).
RN [3]
RN [3] SEQUENCE OF 1-34 FROM N.A.
RA MEDLINE=96096536; PubMed=8522191;
RA Luzi P., Anneschi M., Strayer D.S.;
RT "The upstream region of the SP-B gene: intrinsic promoter activity
RT and glucocorticoid responsiveness related to a new DNA-binding
RT protein.";
RL Gene 165:285-290(1995).
CC -1- FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE
CC ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-
CC LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES
CC THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS
CC PER METER.
CC -1- SUBUNIT: Homodimer; disulfide-linked.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
CC protein. There are 4 surfactant-associated proteins: 2 collagenous,
CC carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small
CC hydrophobic proteins (SP-B and SP-C).
CC -1- SIMILARITY: Contains 1 saposin A-type domain.
CC -1- SIMILARITY: Contains 3 saposin B-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR -----
DR EMBL; M24901; AAA31466.1; -;
DR EMBL; U17106; AAA67934.1; -;
DR EMBL; S80649; AAD14335.1; -;
DR PIR; A32421; LNRBB.
DR PIR; I46531; I46531.
DR HSSP; P07988; 1DFW.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; Sapa_1.
DR InterPro; IPR008138; Sapa_2.
DR InterPro; IPR008140; Sapa_sub.
DR InterPro; IPR008139; SaposinB.
DR InterPro; IPR008137; Surfactant_B.
DR Pfam; PF02199; SAPA; 1.
DR Pfam; PF05184; Sapa_1; 1.
DR Pfam; PF03489; Sapa_2; 3.
DR ProDom; PD001732; Sapa_sub; 1.
DR ProDom; PD008002; Surfactant_B; 1.
DR SMART; SM00162; SAPA; 1.
DR SMART; SM00118; SAPB; 3.
KW Surface film; Gaseous exchange; Glycoprotein; Repeat.
FT PROPEP 1 184
FT CHAIN 185 263 PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT PROPEP 264 370
FT DOMAIN 29 62 SAPOSIN-LIKE TYPE A.
FT DISULFID 192 261 BY SIMILARITY.
FT DISULFID 195 255 BY SIMILARITY.
FT DISULFID 219 230 BY SIMILARITY.
FT DISULFID 232 232 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 300 300 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 15 15 MISSING (IN REF. 2).
FT CONFLICT 184 184 R -> P (IN REF. 1).
FT CONFLICT 232 232 C -> R (IN REF. 2).
FT CONFLICT 329 355 ELHTPQLSLSRGWDARAICQALGAC -> AAHAPAAEPA
FT VQGLGCPNLPFGEGRV (IN REF. 1).
SQ SEQUENCE 370 AA; 40610 MW; 423047A69B12DCB5 CRC64;
Query Match 10.6%; Score 296.5; DB 1; Length 370;
Best Local Similarity 21.7%; Pred. No. 1.7e-12;
Matches 111; Conservative 65; Mismatches 170; Indels 165; Gaps 18;
Qy 4 LFLASLLGAALAGPVLGKCECTGSAVVCQNVKTASDCGAVKCHLQTVNKNPTVKSLPC 63
Db 12 LLLLTLCGPGTAVWATSPACQPEFWCQSEALQCALGCHLQCEVWGHVGADDL-C 70
Qy 64 DICKDVTAAAGDMLKDNATEEILVYLEKTCMDLKPENMSASCKEIVDSYLPVILDIKG 123
Db 71 QECQDIVNLTKEAIFQDTIRKLEHEDVLPLKLLVPOCHHVLVDVFPPLTITVFQS 130
Qy 124 EMSRPGEVCSALNLCESLQKHLAELNKHQKLESNKIPELDMTEVAVPFMANIPLLLYPD 183
Db 131 QINAKA-ICQHLGQC-----PGSPEPLD-----PLPDKLVLPDLL--G 167
Qy 184 GPRSKPOPKONGDVQCQCIQVMTDITQAVRTNSTFVQALVEHVKECDRPGWADICKN 243
Db 168 ALPAKP-----GPHQTJDS-- 181
Qy 244 YISQYSEIAIQMMHMQPKIEICLVGFCDEVKEMPQTLPAKVASKNVIPALEVEPIK 303
Db 182 -----AQRFPPL-----PL- 191
Qy 304 KHEVPAKSDVYCEVCEPLKVEKTKLIDNNKTEKELD-AFDKMCCKLPKSLSECEQVVD 362
Db 192 -----CWLRTLLKRIQAM-----PGVLAMAVAQVCHVPLVVGICQCLAE 235
Qy 363 TYGSSILSILEEVSPELVCSMLHLS-----GTRLPALTVHTVTPKDGCFCEVCKK 415
Db 236 RYTVILLVLLGHVLPQVGLVLRCSVDISIQVFPPTLEALPGEWLPDPPE-CRLCMSV 294
Qy 416 VGYLDNRNLEKSTKOBILAA-----LEKGCSPFLPDPYQKQCDQFVAEVEPVLILVEM 470


```

Db 295 TQAR-RNISEQTRQAVYHACLSQLDK-----QEQFVELHTPQLLSLRGW 343
Oy 471 DPSVFLKIGACPSAHKPLLTGKICMGPSY 501
Db 344 DARAICQALGACVATLSPL-----QCISPHF 370

RESULT 12
SAP_CAVPO
ID SAP_CAVPO STANDARD; PRT; 81 AA.
AC P20097;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Saposin C (CO-beta-glucosidase) (Glucosylceramidase activator)
DE (Sphingolipid activator protein 2) (SAP-2).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=89066787; PubMed=3198642;
RA Sano A., Radin N.S., Johnson L.L., Tarr G.E.;
RT "The activator protein for glucosylceramide beta-glucosidase from
RT guinea pig liver. Improved isolation method and complete amino acid
RT sequence."
RL J. Biol. Chem. 263:19597-19601(1988).
CC -!- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
CC Saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate.
CC -!- SIMILARITY: Contains 1 saposin B-type domain.
DR PIR; A32026; A32026.
DR HSSP; P42210; IQDM.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 1.
DR SMART; SM00118; SAPB; 1.
KW Glycoprotein; Sphingolipid metabolism.
FT DOMAIN 1 81 SAPOSIN-LIKE TYPE B.
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .).
FT DISULFID 5 78 BY SIMILARITY.
FT DISULFID 8 72 BY SIMILARITY.
FT DISULFID 36 47 BY SIMILARITY.
SQ SEQUENCE 81 AA; 8852 MW; E564CE1FOA292596 CRC64;

Query Match 9.4%; Score 261; DB 1; Length 81;
Best Local Similarity 61.8%; Pred. No. 5.5e-11;
Matches 47; Conservative 18; Mismatches 11; Indels 0; Gaps 0;

Oy 313 VYCVCEFLVKEVTKLIDNNKTEKILDAFKMGSKLPKLSSECEQVVDYTGSSILSIL 372
Db 3 VTKACEVYVVKVWELIDNNTEKEIHALDSVALLPESVSEVCEQVVDYTGSDIALL 62

Oy 373 LEEVSPVLCSMLHLC 388
Db 63 LQEMSPELVCSGLGIC 78

RESULT 13
ASPR_CUCPE
ID ASPR_CUCPE STANDARD; PRT; 513 AA.

```

```

AC 004057;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aspartic proteinase precursor (EC 3.4.23.-).
OS Cucurbita pepo (Vegetable marrow) (Summer squash).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OX eucosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
OX NCBI_TaxID=3663;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97354181; PubMed=9210475;
RA Hiraiwa N., Kondo M., Nishimura M., Hara-Nishimura I.;
RT "An aspartic endopeptidase is involved in the breakdown of
RT propeptides of storage proteins in protein-storage vacuoles of
RT plants."
RL Eur. J. Biochem. 246:133-141(1997).
CC -!- FUNCTION: Involved in the breakdown of propeptides of storage
CC proteins in protein-storage vacuoles.
CC -!- SUBUNIT: Heterodimer of a 32 kDa subunit and a 16 kDa subunit.
CC -!- SUBCELLULAR LOCATION: Vacuolar.
CC -!- SIMILARITY: Belongs to peptidase family A1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB002695; BAA19607.1; -
DR PUR; T09739; T09739.
DR HSSP; P42210; IQDM.
DR MEROPS; A01.020; -.
DR InterPro; IPR001969; Asprotease AS.
DR InterPro; IPR009007; Pept A acid.
DR InterPro; IPR001461; Peptidase_A1.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF00026; asp; 1.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR00792; PEPSIN.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 1.
DR SMART; SM00118; SAPB; 2.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Zymogen; Glycoprotein; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT PROPEP 26 ? POTENTIAL.
FT CHAIN ? 513 ASPARTIC PROTEINASE.
FT DOMAIN 322 420 SPECIFIC TO PLANT ASPARTIC PROTEINASES.
FT ACT_SITE 108 108 BY SIMILARITY.
FT ACT_SITE 294 294 BY SIMILARITY.
FT DISULFID 120 126 BY SIMILARITY.
FT DISULFID 285 289 BY SIMILARITY.
FT DISULFID 432 469 BY SIMILARITY.
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 513 AA; 55855 MW; 083FB7064CE02DC2 CRC64;

Query Match 6.2%; Score 173.5; DB 1; Length 513;
Best Local Similarity 29.0%; Pred. No. 0.00031;
Matches 47; Conservative 31; Mismatches 59; Indels 25; Gaps 7;

Oy 353 LSECEQVVDYTGSSILSILLESVPELVCSMLHLC--GTRLPALTVH-VTPKDG--- 406
Db 320 VSQCKAVVAQYGTQIMDLLSEADPKKISQINLCTFDGTRGVSMGIESVVDENACKSS 379

```

```
Oy 407 -----GFEYCKLVGLDRNLEKSTKQELLALEKSGSFLLDPYQK---QCQDFVAEY 458
Db 380 DSLHDGMCSCVETVMVQNLQKQTKERLINIYNELCDRMPSPMGQSAVDCGO-LSSM 438
Oy 459 EPLVIELVEVMD--PSFVCLKIGACPSAKHPLLGTEKIWG 498
Db 439 PTVSFTIGGKIFDLAEEYILKVGEGVA-----QCISG 472

RESULT 14
ASPR_HORVU
ID ASPR_HORVU STANDARD; PRT; 508 AA.
AC P42210;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phytapsin precursor (EC 3.4.23.40) (Aspartic proteinase).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Kustaa; TISSUE=Embryo;
RX MEDLINE=92111473; PubMed=1722454;
RA Runeberg-Roos P., Toormaekangas K., Oestman A.;
RT "Primary structure of a barley-grain aspartic proteinase. A plant
RT aspartic proteinase resembling mammalian cathepsin D.";
RL Eur. J. Biochem. 202:1021-1027(1991).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=cv. Kustaa; TISSUE=Embryo;
RX MEDLINE=92254717; PubMed=1812727;
RA Toormaekangas K., Runeberg-Roos P., Oestman A., Tilgmann C.,
RA Sarkkinen P., Kervinen J., Mikola L., Kalkkinen N.;
RT "Aspartic proteinase from barley seeds is related to animal cathepsin
RT D.";
RL Adv. Exp. Med. Biol. 306:355-359(1991).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=99335466; PubMed=10406799;
RA Kervinen J., Tobin G.J., Costa J., Waugh D.S., Wlodawer A., Zdanov A.;
RT "Crystal structure of plant aspartic proteinase phytypsin:
RT inactivation and vacuolar targeting.";
RL EMBO J. 18:3947-3955(1999).
CC -! CATALYTIC ACTIVITY: Prefers hydrophobic residues Phe, Val, Ile,
CC Leu, and Ala at P1 and P1', but also cleaves -Phe-|-Asp- and
CC -Asp-|-Asp- bonds in 2S albumin from plants seeds
CC -! SUBUNIT: Heterodimer of two subunits (29 kDa and 11 kDa) processed
CC from the precursor molecule. A large enzyme (32 kDa and 16 kDa) is
CC an intermediate precursor form.
CC -! SUBCELLULAR LOCATION: Vacuolar.
CC -! TISSUE SPECIFICITY: Embryo and leaf.
CC -! SIMILARITY: Belongs to peptidase family A1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56136; CAA39602.1; -
CC PIR; S19697; S19697.
CC PDB; 1QDM; 16-JUL-99.
CC MEROPS; A01.020; -
CC InterPro; IPR001969; Asprotease AS.
CC InterPro; IPR009007; Pept A acid.
CC InterPro; IPR001461; Peptidase_A1.
CC InterPro; IPR007856; SapB_1.
```

```
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF00026; asp; 1.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR00792; PEPSIN.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 1.
DR SMART; SM00118; SAPB; 2.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Zymogen; Glycoprotein; Signal;
KW 3D-structure.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 66
FT CHAIN 67 377 PHYTEPSIN 32 kDa SUBUNIT.
FT CHAIN 67 377 PHYTEPSIN 29 kDa SUBUNIT.
FT CHAIN 378 508 PHYTEPSIN 16 kDa SUBUNIT.
FT CHAIN 422 508 PHYTEPSIN 11 kDa SUBUNIT.
FT DOMAIN 317 420 SPECIFIC TO PLANT ASPARTIC PROTEINASES.
FT SITE 377 378 CLEAVAGE.
FT SITE 421 422 CLEAVAGE.
FT ACT_SITE 102 102
FT ACT_SITE 289 289
FT DISULFID 115 121
FT DISULFID 280 284
FT DISULFID 427 464
FT CARBOHYD 399 399
FT STRAND 32 37
FT HELIX 42 49
FT TURN 50 50
FT HELIX 74 76
FT STRAND 79 81
FT TURN 84 90
FT TURN 91 94
FT STRAND 95 102
FT TURN 103 104
FT STRAND 109 112
FT TURN 113 114
FT STRAND 119 123
FT STRAND 127 127
FT HELIX 129 131
FT TURN 133 134
FT STRAND 136 136
FT STRAND 141 146
FT TURN 147 148
FT STRAND 149 162
FT TURN 163 164
FT STRAND 165 177
FT STRAND 180 180
FT HELIX 182 185
FT STRAND 190 193
FT HELIX 197 199
FT HELIX 201 203
FT HELIX 207 211
FT TURN 212 215
FT STRAND 221 225
FT STRAND 236 240
FT TURN 241 241
FT TURN 245 246
FT STRAND 248 260
FT TURN 261 262
FT STRAND 263 267
FT STRAND 270 272
FT TURN 273 274
FT STRAND 275 276
FT TURN 279 282
FT STRAND 284 288
FT STRAND 295 297
FT HELIX 299 309
FT TURN 310 310
FT STRAND 314 314
```

N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT HELIX 317 325
FT TURN 326 326
FT HELIX 327 335
FT TURN 336 337
FT HELIX 340 346
FT TURN 347 348
FT HELIX 379 397
FT TURN 398 399
FT HELIX 402 412
FT TURN 413 414
FT STRAND 416 416
FT STRAND 424 425
FT HELIX 427 430
FT TURN 431 432
FT STRAND 436 440
FT TURN 441 442
FT STRAND 443 447
FT HELIX 449 452
FT STRAND 453 455
FT HELIX 459 461
FT STRAND 464 466
FT STRAND 468 470
FT TURN 475 476
FT STRAND 480 483
FT HELIX 485 488
FT TURN 489 490
FT STRAND 491 496
FT TURN 497 500
FT STRAND 501 508
SQ SEQUENCE 508 AA; 54226 MW; 87F2C9F3369B962 CRC64;

Query Match 5.7%; Score 158.5; DB 1; Length 508;
Best Local Similarity 24.3%; Pred. No. 0.003;
Matches 49; Conservative 40; Mismatches 78; Indels 35; Gaps 9;

Qy 193 DNGDV-----CQDCIQMTDIQAVRTNSTFVALVHVKEBCDRIGPMADICKN 243
Db 266 DMGDLVGGKSTGFCAGCAAIADSGTSLLAGPT---AIITEINEKIGAAGV-VSQECKT 321
Qy 244 YISQYSEIAIQMMW-HMQPKIEICALVGCDEVKEMPMQTLVPAKVASKNVPALVELVEPI 302
Db 322 IVSYGGQILDLLAETQPKKICQVGLC-----TFDGTGVSAGIRSVVD-DEPV 371
Qy 303 KHBEVPAKSDYVCBVFKEVTKLIDNNKTEKILDAFDKMGSKLPKSLSE---EC-- 357
Db 372 KSNGL--RADPMCSACEMAVVMQNLQAQNKTDQLILDYVYNQLCNRLSPMGESAVDCGS 429
Qy 358 -----QEVVDYVGSSILSILEE 375
Db 430 LGSMPDIEFTIGGKFKPKPEE 451

```

RESULT 15

```

CYP1_CYNCA STANDARD; PRT; 473 AA.
AC P40782;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cyprosin precursor (EC 3.4.23.-) (Fragment).
GN CYP101.
OS Cynara cardunculus (Cardoon).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Carduoideae; Cardueae; Cynara.
OK NCBI_TaxID=4265;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 178-186.
RC TISSUE=Flower buds;
RX MEDLINE=94250836; PubMed=8193298;
RA Cordero M.C.; Xue Z.-T.; Pietrzak M.; Pais M.S.; Brodelius P.E.;
RT "Isolation and characterization of a cDNA from flowers of Cynara
cardunculus encoding cyprosin (an aspartic proteinase) and its use to

```

```

RT study the organ-specific expression of cyprosin.";
RL Plant Mol. Biol. 24:733-741(1994).
CC -I- TISSUE SPECIFICITY: Mostly present in the violet parts of styles
CC and corollas of mature flowers.
CC -I- DEVELOPMENTAL STAGE: Expressed in early stages of floral
CC development and switched off at maturation of the flower.
CC -I- SIMILARITY: Belongs to peptidase family A1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; X69193; CAA48939.1; ALT_SEQ.
DR HSSP; P42210; IQDM.
DR MEROPS; A01.020; -.
DR InterPro; IPR001969; Aspartic protease AS.
DR InterPro; IPR009007; Peptidase A1.
DR InterPro; IPR001461; Peptidase A1.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB sub.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF00026; asp_1.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR00792; PEPFIS.
DR ProDom; PD001732; SapB sub; 1.
DR SMART; SM00118; SAPB; 2.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Zymogen; Glycoprotein.
FT NON_TER 1
FT PROPEP <1 33 POTENTIAL.
FT CHAIN 34 473 CYPROSIN.
FT ACT_SITE 69 69 BY SIMILARITY.
FT ACT_SITE 256 256 BY SIMILARITY.
FT DISULFID 82 88 POTENTIAL.
FT DISULFID 247 251 POTENTIAL.
FT DISULFID 392 429 POTENTIAL.
FT CARBOHYD 364 364 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 473 AA; 51564 MW; 65F3232EBD06CB56 CRC64;

Query Match 5.5%; Score 154.5; DB 1; Length 473;
Best Local Similarity 30.1%; Pred. No. 0.005;
Matches 31; Conservative 24; Mismatches 39; Indels 9; Gaps 2;

Qy 353 LSEECQEVVDYTGSSILSILEEVSPELVCSMLHLC--GTRLPALTVHVTQPKDGG--- 407
Db 282 MSQCKSLVDQYQKSMIEMLLSEQPEKICQMKLCSFDGSHDTSMLIESVVDKSKSS 341
Qy 408 ----FCEVCKKLYGLDRNLKRNKSTKQETLAALAEKGCSPFLDP 446
Db 342 GLPMRCVPCARWVMQNIQRQNETENIINYVDKLCERLSP 384

Search completed: May 5, 2004, 13:31:42
Job time : 14.9383 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:26:23 ; Search time 20.9003 Seconds
(without alignments)
2411.659 Million cell updates/sec

Title: US-09-743-684A-1
Perfect score: 2789
Sequence: 1 MYALFLLASLLGALAGPVL.....NTETAQCNAVECHKRHVMN 524

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2777.5	99.6	527	1 SAHUP	saposin precursor
2	1996	71.6	554	1 A28716	saposin precursor
3	1937.5	69.5	557	1 JH0604	saposin precursor
4	511	18.3	965	2 T00207	P109 protein - sil
5	389	13.9	79	2 A49475	cerebroside sulfat
6	379	13.6	376	2 S02766	pulmonary surfacta
7	348.5	12.5	381	1 LNHUB	pulmonary surfacta
8	343	12.3	80	2 S21770	saposin-C - bovine
9	310	11.1	363	2 A29072	pulmonary surfacta
10	301.5	10.8	369	2 I46531	surfactant protein
11	261	9.4	81	2 A32036	glucosylceramide b
12	256.5	9.2	370	1 LNRBB	pulmonary surfacta
13	213	7.6	213	2 T46069	hypothetical prote
14	203.5	7.3	217	2 T48201	hypothetical prote
15	176.5	6.3	402	2 T15677	hypothetical prote
16	173.5	6.2	513	2 T09739	hypothetical prote
17	171.5	6.1	506	2 F86253	hypothetical prote
18	165.5	5.9	506	2 T07915	probable aspartic
19	161.5	5.8	513	2 T11686	aspartic proteinas
20	161	5.8	428	2 S47096	cynarase (EC 3.4.2
21	161	5.8	474	2 T12049	cyprosin (EC 3.4.2
22	158.5	5.7	508	2 T19697	aspartic proteinas
23	156	5.6	314	2 T15674	hypothetical prote
24	155.5	5.6	292	2 T14446	aspartic proteinas
25	155.5	5.6	322	2 S41400	aspartic proteinas
26	154	5.5	496	2 JS0732	aspartic proteinas
27	153.5	5.5	509	2 JC7272	aspartic proteinas
28	145.5	5.2	205	2 B89567	protein T08A9.7 [i
29	143.5	5.1	1175	2 S52417	E-selectin ligand-

ALIGNMENTS

RESULT 1
SAHUP

saposin precursor [validated] - human

N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component ein (SAP); sphingolipid activator protein A2; sulfatide sulfatase activator protein N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1992 #sequence revision 17-Nov-1995 #text change 08-Dec-2000
C;Accession: JX0061; A57368; A42003; B42003; C42003; D42003; A30367; S34740; S36140; S36
0226; I37265; I37264
R;Nakano, T.; Sandhoff, K.; Stuenkel, J.; Christomanou, H.; Suzuki, K.
J. Biochem. 105, 152-154, 1989
A;Title: Structure of full-length cDNA coding for sulfatide activator, a Co-beta-glucosidase
A;Reference number: JX0061; MUID:89255151; PMID:2498298
A;Accession: JX0061
A;Molecule type: mRNA
A;Residues: 1-527 <NAK>
A;Cross-references: GB:D00422; NID:g220063; PIDN:BAA00321.1; PID:g220064
A;Note: alternative splice form 1
A;Accession: A57368
A;Molecule type: mRNA
A;Residues: 1-259,263-527 <NA2>
A;Cross-references: GB:J03015; GB:J03086; NID:G337755; PIDN:AAB59494.1; PID:G337756
A;Note: alternative splice form 2
R;Rorman, E.G.; Scheinker, V.; Grabowski, G.A.
Genomics 13, 312-318, 1992
A;Title: Structure and evolution of the human prosaposin chromosomal gene.
A;Reference number: A42003; MUID:92307663; PMID:1612590
A;Accession: A42003
A;Molecule type: DNA
A;Residues: 50-140 <ROR>
A;Cross-references: GB:M86181
A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107236)
A;Accession: B42003
A;Molecule type: DNA
A;Residues: 185-259,263-276 <RO2>
A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107237)
A;Accession: C42003
A;Molecule type: DNA
A;Residues: 305-393 <RO3>
A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107238); sequence inc
A;Accession: D42003
A;Molecule type: DNA
A;Residues: 399-487 <RO4>
A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107239); sequence inc
R;Rorman, E.G.; Grabowski, G.A.
Genomics 5, 486-492, 1989
A;Title: Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four spin
A;Reference number: A30367; MUID:90129043; PMID:2515150
A;Accession: A30367
A;Molecule type: mRNA
A;Residues: 1-259,263-527 <RO5>

oryzasin (EC 3.4.2
cyprosin (EC 3.4.2
gene 11-1 protein
probable aspartic
cag pathogenicity
hypothetical prote
aspartic proteinas
cysteine-rich fibr
hypothetical prote
cag island protein
aspartic proteinas
hypothetical prote
probable CHP-rich
hypothetical prote
rsec8 - rat (fragm

A;Cross-references: GB:J03077; NID:gi83230; PIDN:AAA52560.1; PID:g183231
A;Note: alternative splice form 2
R;Hiraiwa, M.; O'Brien, J.S.; Kishimoto, Y.; Galdsicka, M.; Fluharty, A.L.; Ginns, E.I.; Arch. Biochem. Biophys. 304, 110-116, 1993
A;Title: Isolation, characterization, and proteolysis of human prosaposin, the precursor of saposin A
A;Reference number: S34740; MUID:93311991; PMID:8323276
A;Accession: S34740
A;Molecule type: protein
A;Residues: 17-24;165-172;180-189;301-305 <HIR>
R;Tyynelae, J.; Palmer, D.N.; Baumann, M.; Haltia, M. FEBS Lett. 330, 8-12, 1993
A;Title: Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis.
A;Reference number: S36140; MUID:93380576; PMID:8370464
A;Accession: S36140
A;Molecule type: protein
A;Residues: 'XX', 62, 'X', 64-65, 'X', 67-79, 'X', 81-84 <TY>
A;Note: saposin A
A;Accession: S36141
A;Molecule type: protein
A;Residues: 'XX', 413-414, 'X', 416-428, 'X', 430-434 <TV2>
A;Note: saposin D
R;Holtzman, H.; Sandhoff, K.; Kwon, H.Y.; Harzer, K.; Nakano, T.; Suzuki, K. J. Biol. Chem. 266, 7556-7560, 1991
A;Title: Sulfatide activator protein. Alternative splicing that generates three mRNAs and a protein
A;Reference number: S36988; MUID:91210267; PMID:2019586
A;Accession: S36988
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-240, 'S', 242-527 <HOL>
A;Cross-references: EMBL:M60255; NID:g337759; PIDN:AAA36594.1; PID:g337760
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
A;Note: cerebrosulfate activator protein mutant MU-9; corresponds to alternative splicing
A;Accession: S36989
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-240, 'S', 242-259, 263-527 <HOL>
A;Cross-references: EMBL:M60257; NID:g337764; PIDN:AAA36595.1; PID:g337765
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
A;Note: cerebrosulfate activator protein mutant MU-0; corresponds to alternative splicing
A;Accession: S36990
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-240, 'S', 242-259, 261-527 <HOL>
A;Cross-references: EMBL:M60258; NID:g337766; PIDN:AAA36596.1; PID:g337767
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
A;Note: cerebrosulfate activator protein mutant MU-6; corresponds to alternative splicing
R;Kondoh, K.; Hinenio, T.; Sano, A.; Kakimoto, Y. Biochem. Biophys. Res. Commun. 181, 286-292, 1991
A;Title: Isolation and characterization of prosaposin from human milk.
A;Reference number: PS0330; MUID:92068206; PMID:1958198
A;Accession: PS0330
A;Molecule type: protein
A;Residues: 17-24, 'X', 26 <KON>
A;Experimental source: milk
R;Kretz, K.A.; Carson, G.S.; Morimoto, S.; Kishimoto, Y.; Fluharty, A.L.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 87, 2541-2544, 1990
A;Title: Characterization of a mutation in a family with saposin B deficiency: a glycosylation defect
A;Reference number: A35985; MUID:90207231; PMID:2320574
A;Accession: A35985
A;Molecule type: mRNA
A;Residues: 213-221 <KRE>
A;Cross-references: GB:M32221; NID:g337761; PIDN:AAA60303.1; PID:g337762
A;Accession: B35985
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-259, 263-527 <KR2>
A;Cross-references: GB:M32221; NID:g337761; PIDN:AAA60303.1; PID:g337762
A;Experimental source: lymphoblast
A;Accession: C35985
A;Molecule type: mRNA
A;Residues: 213-216, 'I', 218-221 <KR3>
A;Note: sequence from patients with activator-deficient metachromatic leukodystrophy; the sequence was revised relative to PID:g337761
R;Fuerst, W.; Schubert, J.; Machleidt, W.; Meyer, H.E.; Sandhoff, K.

Eur. J. Biochem. 192, 709-714, 1990
A;Title: The complete amino-acid sequences of human ganglioside GM2 activator protein and its precursor
A;Reference number: S13195; MUID:91006185; PMID:2209618
A;Accession: S13196
A;Molecule type: protein
A;Residues: 195-259, 263-277 <FUE>
R;Morimoto, S.; Martin, B.M.; Yamamoto, Y.; Kretz, K.A.; O'Brien, J.S.; Kishimoto, Y. Proc. Natl. Acad. Sci. U.S.A. 86, 3389-3393, 1989
A;Title: Saposin A: second cerebrosulfate activator protein.
A;Reference number: A32784; MUID:89240739; PMID:2717620
A;Accession: A32784
A;Molecule type: protein
A;Residues: 60-84; 86-107; 109-119; 125-134 <MOR>
R;O'Brien, J.S.; Kretz, K.A.; Dewji, N.; Wenger, D.A.; Esch, F.; Fluharty, A.L. Science 241, 1098-1101, 1988
A;Title: Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic information
A;Reference number: A41240; MUID:88321660; PMID:2842863
A;Accession: A41240
A;Molecule type: mRNA
A;Residues: 'GSSR', 18-259, 263-299, 'D', 301-302, 'D', 304-527 <OAB>
A;Cross-references: GB:J03086
R;Dewji, N.N.; Wenger, D.A.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 84, 8652-8656, 1987
A;Title: Nucleotide sequence of cloned cDNA for human sphingolipid activator protein 1 p1
A;Reference number: S02289; MUID:88068647; PMID:2825202
A;Accession: S02289
A;Status: significant sequence differences
A;Molecule type: mRNA
A;Cross-references: EMBL:J03015
A;Note: this sequence corrected by A41240
R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 369, 1361-1365, 1988
A;Title: Complete amino-acid sequence of the naturally occurring A(2) activator protein 1 p1
A;Reference number: S02028; MUID:89207118; PMID:3242555
A;Accession: S02028
A;Molecule type: protein
A;Residues: 195-259, 263-276 <KLE>
R;Fuerst, W.; Machleidt, W.; Sandhoff, K. Biol. Chem. Hoppe-Seyler 369, 317-328, 1988
A;Title: The precursor of sulfatide activator protein is processed to three different products
A;Reference number: S00813; MUID:89000190; PMID:3048308
A;Accession: S00813
A;Molecule type: protein
A;Residues: 410-487 <PU2>
R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 368, 1571-1578, 1987
A;Title: Complete amino-acid sequence and carbohydrate content of the naturally occurring A(2) activator protein 1 p1
A;Reference number: S00226; MUID:88163077; PMID:3442600
A;Accession: S00226
A;Molecule type: protein
A;Residues: 314-393 <KL2>
R;Vacarro, A.M.; Salvio, R.; Barca, A.; Tatti, M.; Ciaffoni, F.; Maras, B.; Siciliano, J. Biol. Chem. 270, 9953-9960, 1995
A;Title: Structural analysis of saposin C and B. Complete localization of disulfide bridge
A;Reference number: A57297; MUID:95247790; PMID:7730378
R;Contents: annotation; disulfide bonds; glycosylation
R;Holtzman, H.; Sandhoff, K.; Fuerst, W.; Kwon, H.Y.; Schnabel, D.; Suzuki, K. FEBS Lett. 280, 267-270, 1991
A;Title: The organization of the gene for the human cerebrosulfate sulfate activator protein
A;Reference number: I37264; MUID:91192146; PMID:2013321
A;Accession: I37265
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 59-125 <RES>
A;Cross-references: EMBL:X57107; NID:g30234; PIDN:CAA40391.1; PID:g30235
A;Accession: I37264
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 307-516 <RE2>
A;Cross-references: EMBL:X57108; NID:g30232; PIDN:CAA40392.1; PID:g1565257
A;Note: sequence revised relative to PID:g30233 (corrected coding region)
C;Genetics:

A;Gene: GDB:PSAP; GLBA
A;Cross-references: GDB:120366; OMIM:176801
A;Map position: 10q22.1-10q22.1
A;Introns: 83/3; 338/3; 401/1; 453/3; 480/3
A;Note: defects in this gene may cause variant Gaucher disease, variant Tay-Sachs disease
A;Note: list of introns is incomplete

Query Match 99.6%; Score 2777.5; DB 1; Length 527;
Best Local Similarity 99.4%; Pred. No. 1.1e-175;
Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 MYALFLASLLGAALAGPVGLGKCTRGSAVWCNQVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 1 MYALFLASLLGAALAGPVGLGKCTRGSAVWCNQVKTASDCGAVKHCLQTVWVKPTVKS 60
Qy 61 LPCDICVTVTAAGDMLKDNATEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICVTVTAAGDMLKDNATEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLY 180
Qy 181 PDGPRSKPOPKONGDVCDQCIOQVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGWADI 240
Db 181 PDGPRSKPOPKONGDVCDQCIOQVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGWADI 240
Qy 241 CKNYISQYSEIAIOMMHMOPKEICALVGFCDVEKEMPOTLVPAKVASKNVIIPALE 297
Db 241 CKNYISQYSEIAIOMMHMOPKEICALVGFCDVEKEMPOTLVPAKVASKNVIIPALE 300
Qy 298 LVEPIKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEEC 357
Db 301 LVEPIKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEEC 360
Qy 358 QEVVDTYGSSTLSLLEVSPELVCSMLHLCSTGRLPALTVHVTQPKDGGFCEVCKKLVG 417
Db 361 QEVVDTYGSSTLSLLEVSPELVCSMLHLCSTGRLPALTVHVTQPKDGGFCEVCKKLVG 420
Qy 418 YLDRLNLEKNSTKQIBILAALKEGCSFLDPYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCL 477
Db 421 YLDRLNLEKNSTKQIBILAALKEGCSFLDPYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCL 480
Qy 478 KIGACPSAHPKLLGTEKCIWGPSYWCQNTETAACNAVEHCKRHVN 524
Db 481 KIGACPSAHPKLLGTEKCIWGPSYWCQNTETAACNAVEHCKRHVN 527

RESULT 2

A28716
saposin precursor - rat
N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfata
N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A28716
R;Collard, M.W.; Sylvestre, S.R.; Tsuruta, J.K.; Griswold, M.D.
Biochemistry 27, 4557-4564, 1988
A;Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat s
A;Reference number: A28716; MUID:89000647; PMID:3048385
A;Accession: A28716
A;Molecule type: mRNA
A;Residues: 1-554 <COL>
A;Cross-references: GB:M19936; NID:g206904; PIDN:AAA42136.1; PID:g206905
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
C;Function:
A;Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc
A;Pathway: sphingolipid catabolism
A;Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy
A;Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsulfi
A;Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiester
C;Superfamily: saposin; saposin repeat homology

C;Keywords: alternative splicing; glycoprotein; lysoproteins; lysosomal storage disease; sph
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-554/Product: prosaposin #status predicted <PRO>
F;55-148/Domain: saposin repeat homology <SAP1>
F;60-143/Product: saposin A #status predicted <SAPA>
F;189-280/Domain: saposin repeat homology <SAP2>
F;194-273/Product: saposin B #status predicted <SAB1>
F;306-397/Domain: saposin repeat homology <SAP3>
F;310-389/Product: saposin C #status predicted <SAPC>
F;431-522/Domain: saposin repeat homology <SAP4>
F;437-514/Product: saposin D #status predicted <SAPD>
F;63-138,66-132,94-106,439-512,442-506,470-481/Disulfide bonds: #status predicted
F;80,214,331,456/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;197-270,200-264,229-240,314-387,317-381,345-356/Disulfide bonds: #status predicted

Query Match 71.6%; Score 1996; DB 1; Length 554;
Best Local Similarity 66.3%; Pred. No. 4.1e-124;
Matches 368; Conservative 77; Mismatches 78; Indels 32; Gaps 3;

Qy 1 MYALFLASLLGAALAGPVGLGKCTRGSAVWCNQVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 1 MYALFLASLLGAALAGPVGLGKCTRGSAVWCNQVKTASDCGAVKHCLQTVWVKPTVKS 60
Qy 61 LPCDICVTVTAAGDMLKDNATEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICVTVTAAGDMLKDNATEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLY 179
Qy 181 PDGPRSKPOPKONGDVCDQCIOQVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGWADI 240
Db 180 PDGPRSKPOPKONGDVCDQCIOQVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGWADI 239
Qy 241 CKNYISQYSEIAIOMMHMOPKEICALVGFCDVEKEMPOTLVPAKVASKNVIIPALEIVE 300
Db 240 CKNYISQYSEIAIOMMHMOPKEICALVGFCDVEKEMPOTLVPAKVASKNVIIPALEIVE 299
Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEECQEV 360
Db 300 PYQDVIQANVIFCQCVLMVKLSLIIINNTEELIKGLSKACSLPAPATKRCQEV 359
Qy 361 VDTYGSSTLSLLEVSPELVCSMLHLCSTGRLPALTVHVTQPKDGGFCEVCKKLVG 394
Db 360 LVTFGPSLLDVLMEVNPFLCGVISLCSANPNLVGTLEQPAALVSAIPKEAPPKQPE 419
Qy 395 ----ALTAVHTQPKDGGFCEVCKKLVGYLDRLNLEKNSTKQIBILAALKEGCSFLDPYQK 449
Db 420 EPKQSAALRAHVPQKNGGFEVCKKLVYILEHNLKNSTKEEILAALEKGCSPDPYQK 479
Qy 450 QCDOFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAHPKLLGTEKCIWGPSYWCQNTETA 509
Db 480 QCDOFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAHPKLLGTEKCIWGPSYWCQNTETA 539
Qy 510 AQCNVAHEHCKRHVN 524
Db 540 ARCNVADHCKRHVN 554

RESULT 3
JH0604

saposin precursor - mouse
N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfata
N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: JH0604
R;Tsuada, M.; Sakiyama, T.; Endo, H.; Kitagawa, T.
Biochem. Biophys. Res. Commun. 184, 1266-1272, 1992
A;Title: The primary structure of mouse saposin.
A;Reference number: JH0604; MUID:92272718; PMID:1590788

A;Accession: JH0604
A;Molecule type: mRNA
A;Residues: 1-557 <TSU>
A;Cross-references: GB:536200; NID:G249386; PIDN:AAB22175.1; PID:G249387
A;Experimental source: liver
C;Function:
A;Description: saposins bind sphingolipids, form hydrophilic complexes and make them accessible for degradation.
A;Pathway: sphingolipid catabolism
A;Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosylase.
A;Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsulfatase.
A;Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiesterase.
A;Superfamily: saposin; saposin repeat homology
C;Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sphingolipid
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-557/Product: prosaposin #status predicted <PRO>
F;55-148/Domain: saposin repeat homology <SAP1>
F;55-143/Product: saposin A #status predicted <SAP1>
F;189-283/Domain: saposin repeat homology <SAP2>
F;194-276/Product: saposin B #status predicted <SAB1>
F;309-400/Domain: saposin repeat homology <SAP3>
F;313-392/Product: saposin C #status predicted <SAPC>
F;434-525/Domain: saposin repeat homology <SAP4>
F;440-517/Product: saposin D #status predicted <SAPD>
F;63-138,66-132,94-106,197-273,200-267,229-240,317-390,320-384,348-359,442-515,445-509,450-515,517-557/Binding site: carbohydrate (Asn) #status predicted
Query Match 69.5%; Score 1937.5; DB 1; Length 557;
Best Local Similarity 63.6%; Pred. No. 3e-120;
Matches 355; Conservative 77; Mismatches 91; Indels 35; Gaps 3;
QY 1 MYALFLLASLGAALAGPVLGKETRGSAVWQNVKTSASDCAVGHCLQTVWVKPTVKS 60
Db 1 MYALFLLASLGAALAGPVLGKETRGSAVWQNVKTSASDCAVGHCLQTVWVKPTVKS 60
QY 61 LPDICKDVTAAAGDMLKDNATEEILVLEKTCMDLKPKNMSASCKEIVDSYLPVILDI 120
Db 61 LPDICKDVTAAAGDMLKDNATEEILVLEKTCMDLKPKNMSASCKEIVDSYLPVILDI 120
QY 121 IKGNSRSGVCSALNLCESLOKHLAEHLNKHQLESNKIPELDMTEVVAPEMANIPLLY 180
Db 121 IKGNSRSGVCSALNLCESLOKHLAEHLNKHQLESNKIPELDMTEVVAPEMANIPLLY 180
QY 181 PQDGRSPKPKDNGDVCCDQCIQWVTDIOTAVRTNSTFVQALVHVKECDRLGPGMADI 240
Db 181 PQDGRSPKPKDNGDVCCDQCIQWVTDIOTAVRTNSTFVQALVHVKECDRLGPGMADI 240
QY 241 CKNYISQYSEIAIQMMHMM---QPKKEICALVGFCDVEKMPMOTLVPAKVASKNVIPALE 297
Db 241 CKNYISQYSEIAIQMMHMM---QPKKEICALVGFCDVEKMPMOTLVPAKVASKNVIPALE 297
QY 298 LVEPIKXHEVPAKSDVYCEFLVKEVTKLIDNNKTEKILDAFDMCKSKLPKSLSEEC 357
Db 298 LVEPIKXHEVPAKSDVYCEFLVKEVTKLIDNNKTEKILDAFDMCKSKLPKSLSEEC 357
QY 300 MMDPYEQNLVQAHNVILCQTCQFVWVKFSELI VNNATELLVKGSLNACGVLPDPARTKC 359
Db 300 MMDPYEQNLVQAHNVILCQTCQFVWVKFSELI VNNATELLVKGSLNACGVLPDPARTKC 359
QY 358 QEVVDYTGSSILSILLESVPELVCSMLHLCGSG----- 390
Db 358 QEVVDYTGSSILSILLESVPELVCSMLHLCGSG----- 390
QY 360 QEVVDYTGSSILSILLESVPELVCSMLHLCGSG----- 390
Db 360 QEVVDYTGSSILSILLESVPELVCSMLHLCGSG----- 390
QY 391 ----TRLPALTVMVTPQDGGFCVCKLVGYLDRLNKNSTKQETLAALAEKGCSPFLDP 419
Db 391 ----TRLPALTVMVTPQDGGFCVCKLVGYLDRLNKNSTKQETLAALAEKGCSPFLDP 419
QY 420 QPAQPKOSALPAHVPPQKNGGFCVCKLVGYLDRLNKNSTKQETLAALAEKGCSPFLDP 479
Db 420 QPAQPKOSALPAHVPPQKNGGFCVCKLVGYLDRLNKNSTKQETLAALAEKGCSPFLDP 479
QY 447 YQKQCDQFVAEYEPFLVLEILVEVMDPSFVCLTKGACPSAHKPLLGTEKIMGVSWCQNT 506
Db 447 YQKQCDQFVAEYEPFLVLEILVEVMDPSFVCLTKGACPSAHKPLLGTEKIMGVSWCQNT 506
QY 480 YQKQCDQFVAEYEPFLVLEILVEVMDPSFVCLTKGACPSAHKPLLGTEKIMGVSWCQNT 539
Db 480 YQKQCDQFVAEYEPFLVLEILVEVMDPSFVCLTKGACPSAHKPLLGTEKIMGVSWCQNT 539
QY 507 ETAAQCNVACHCKRHWN 524
Db 507 ETAAQCNVACHCKRHWN 524
QY 540 ETAAQCNVACHCKRHWN 557
Db 540 ETAAQCNVACHCKRHWN 557

RESULT 4
T00207

P109 protein - silkworm
C;Species: Bombyx mori (silkworm)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C;Accession: T00207
R;Ambunan, J.; Chang, P.K.; Li, H.; Natori, M.
Gene 212, 287-293, 1998
A;Title: Molecular cloning of a cDNA encoding a silkworm protein which contains the conserved domain of saposin.
A;Accession: T00207
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-965 <TAM>
A;Cross-references: EMBL:AB008449; NID:G2575864; PIDN:BAA23126.1; PID:G2575865
C;Superfamily: saposin repeat homology
F;778-870/Domain: saposin repeat homology <SAP3>
Query Match 18.3%; Score 511; DB 2; Length 965;
Best Local Similarity 24.6%; Pred. No. 7.4e-26;
Matches 139; Conservative 104; Mismatches 229; Indels 94; Gaps 18;
QY 19 VLGLKECTRGSAVWQNVKTSASDCAVGHCLQTVWVK---PTVKSILPCDICKDVVTAAGD 75
Db 19 VLGLKECTRGSAVWQNVKTSASDCAVGHCLQTVWVK---PTVKSILPCDICKDVVTAAGD 75
QY 170 LLGSKRCTWGPSVWCSNFSSTGRECNATPHCINRWWSKMTFFPDNDNICQICLDMVKQARD 229
Db 170 LLGSKRCTWGPSVWCSNFSSTGRECNATPHCINRWWSKMTFFPDNDNICQICLDMVKQARD 229
QY 76 MLKDNATEEILVLEKTCMDLKPKNMSASCKEIVDSYLPVILDIKGMRSRPGSVCSAL 135
Db 76 MLKDNATEEILVLEKTCMDLKPKNMSASCKEIVDSYLPVILDIKGMRSRPGSVCSAL 135
QY 230 QLOSNBTODEIKEVPEGSCKLPIKFAVAGCKLADEFVVELIETLASEMN-PQAVCSVA 288
Db 230 QLOSNBTODEIKEVPEGSCKLPIKFAVAGCKLADEFVVELIETLASEMN-PQAVCSVA 288
QY 136 NLCESS--LOKHLAEHLNKHQLES-----NKIPELDMTEVVAPEMANIPLLY 173
Db 136 NLCESS--LOKHLAEHLNKHQLES-----NKIPELDMTEVVAPEMANIPLLY 173
QY 289 GLCNNAKIDRLVDVNAQRELKAGCYNQKTVGVVRRKFDTEKDYEDFLVGLLQVCRNMD 348
Db 289 GLCNNAKIDRLVDVNAQRELKAGCYNQKTVGVVRRKFDTEKDYEDFLVGLLQVCRNMD 348
QY 174 ---NIPLLLYP-----QDG-----PRSKPQPKDNGD 196
Db 174 ---NIPLLLYP-----QDG-----PRSKPQPKDNGD 196
QY 349 LSDSCSMLIFKYENILEAVKDLNPEGI CHVSGCYSYKFNHDEFTFPEQMVQYSATDD 408
Db 349 LSDSCSMLIFKYENILEAVKDLNPEGI CHVSGCYSYKFNHDEFTFPEQMVQYSATDD 408
QY 197 V-CODCIQWVTDIOTAVRTNSTFVQALVHVKECDRLGPGMADICKNYISQYSEIAIQM 255
Db 197 V-CODCIQWVTDIOTAVRTNSTFVQALVHVKECDRLGPGMADICKNYISQYSEIAIQM 255
QY 409 VPCEFCQGLVHKRLDVLVANTTELE-FYKVLQGLCKQTKG-FKDECLHLABQYYPVINYF 466
Db 409 VPCEFCQGLVHKRLDVLVANTTELE-FYKVLQGLCKQTKG-FKDECLHLABQYYPVINYF 466
QY 256 MM-HMQPKEICALVGFCDVEKMPMOTLV-----PAKVA---SKNVIP 294
Db 256 MM-HMQPKEICALVGFCDVEKMPMOTLV-----PAKVA---SKNVIP 294
QY 467 LVSDLKPAETCKMIGICGNLTSAPISLVARELVVQPKLIGAESKIAKRVPLAKQMEP 526
Db 467 LVSDLKPAETCKMIGICGNLTSAPISLVARELVVQPKLIGAESKIAKRVPLAKQMEP 526
QY 295 ALELVE--PIKXHEVPA-KSDVYCEFLVKEVTKLIDNNKTEKILDAFDMCKSKLPK 351
Db 295 ALELVE--PIKXHEVPA-KSDVYCEFLVKEVTKLIDNNKTEKILDAFDMCKSKLPK 351
QY 527 ASAAVSVLPLERMFVAAPQSKAACAFQCYFLHVLQVLSDTRTEDKVAQAQVACDALPD 586
Db 527 ASAAVSVLPLERMFVAAPQSKAACAFQCYFLHVLQVLSDTRTEDKVAQAQVACDALPD 586
QY 352 SLSEBCEQVVDYTGSSILSILLESVPELVCSMLHLCGSGTRLPALTVMVTPQDGGFCV 411
Db 352 SLSEBCEQVVDYTGSSILSILLESVPELVCSMLHLCGSGTRLPALTVMVTPQDGGFCV 411
QY 587 ALNGECKEFVQYGSVAILLVQEI DPASVCPALQICQTE-EIRRVNVNSEKSN--CPL 643
Db 587 ALNGECKEFVQYGSVAILLVQEI DPASVCPALQICQTE-EIRRVNVNSEKSN--CPL 643
QY 412 CKKLVGYLDRLNKNSTKQETLAALAEKGCSPFLDPYQKQCDQFVAEYEPFLVLEILVEVMD 471
Db 412 CKKLVGYLDRLNKNSTKQETLAALAEKGCSPFLDPYQKQCDQFVAEYEPFLVLEILVEVMD 471
QY 644 CLFAVEQLESVLKNNRSEENIRKALDGLCTRLSQKQSECIDFVDVTVSSQLVEMLVADM 703
Db 644 CLFAVEQLESVLKNNRSEENIRKALDGLCTRLSQKQSECIDFVDVTVSSQLVEMLVADM 703
QY 472 PSFVCLKIGAC-PSAHKPLLGTEKCI 496
Db 472 PSFVCLKIGAC-PSAHKPLLGTEKCI 496
QY 704 AKEICVFLKLCRQLDHPDLKTHSSI 729
Db 704 AKEICVFLKLCRQLDHPDLKTHSSI 729
RESULT 5
A49475
cerebroside sulfate activator - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 17-May-1996
C;Accession: A49475
R;Stevens, R.L.; Faull, K.F.; Conklin, K.A.; Green, B.N.; Fluharty, A.L.
Biochemistry 32, 4051-4059, 1993
A;Title: Porcine cerebroside sulfate activator: further structural characterization and cDNA cloning.
A;Reference number: A49475; MUID:93229506; PMID:8471613
A;Accession: A49475
A;Status: preliminary

A:Molecule type: protein
A:Residues: 1-79 <STE>
A:Experimental source: kidney
A>Note: sequence extracted from NCBI backbone (NCBIP:129597)
C:Superfamily: saposin; saposin repeat homology
F:1-79/Domain: saposin repeat homology <SAP>

Query Match 13.9%; Score 389; DB 2; Length 79;
Best Local Similarity 88.6%; Pred. No. 3.6e-19;
Matches 70; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 195 GDVQCDCIQWTDIQTAVRTNSTFVQALVHVHKECDRLGPGMADICKNVIQYSEIAIQ 254
Db 1 GDVQCDCIQWTDIQTAVRTNSTFVQALVHVHKECDRLGPGMADICKNVIQYSEIAIQ 60
Qy 255 MMHMQPKEICALVGFCD 273
Db 61 MMHMQPKEICALVGFCEE 79

RESULT 6
S02766
pulmonary surfactant protein B precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 20-Aug-1999
C:Accession: S02766
R:Emrie, P.A.; Shannon, J.M.; Mason, R.J.; Fisher, J.H.
Biochim. Biophys. Acta 994, 215-221, 1989
A>Title: cDNA and deduced amino acid sequence for the rat hydrophobic pulmonary surfactant protein B precursor
A:Reference number: S02766; MUID:89150284; PMID:2920185
A:Accession: S02766
A:Molecule type: mRNA
A:Residues: 1-376 <EMR>
A:Cross-references: EMBL:X14778; NID:957284; PIDN:CAA32885.1; PID:G57285
C:Superfamily: pulmonary surfactant protein B; saposin repeat homology
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-190/Domain: propeptide #status predicted <PRO>
F:59-151/Domain: saposin repeat homology <SAP1>
F:190-277/Domain: saposin repeat homology <SAP2>
F:191-269/Product: pulmonary surfactant protein B #status predicted <MAT>
F:286-371/Domain: saposin repeat homology <SAP3>

Query Match 13.6%; Score 379; DB 2; Length 376;
Best Local Similarity 22.6%; Pred. No. 1.2e-17;
Matches 116; Conservative 75; Mismatches 158; Indels 164; Gaps 16;

Qy 4 LFLIASLGLAAGPLVGLKCTRGSAVQNVKTSDCGAVKICLOTWKNKPTVKSLPC 63
Db 9 LLLPTLCSLGAATESASSPDCAQSPKWCQSLQAIQCRALGHCLQEVMGHAGANDL-C 67
Qy 64 DICDDVVTAAAGMDLKDNTATEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVLDLIKG 123
Db 68 QECEDIHLLTKYTKEDAFQDTIRKFLQECEDILPLKLLVPRCQVLDVLPVLDIVFQG 127
Qy 124 EMSRPGVSGALNICESLQKHLAEHLNQKLESNKIPELDMTEVYAPFMANIPILLYPOD 183
Db 128 QI-KPKAICSHVGLC-----PL--GQT 146
Qy 184 GPRSKPQKNGDVCQDCIQWTDIQTAVRTNSTFVQALVHVHKECDRLGPGMADICKN 243
Db 147 KPEQKE----- 153
Qy 244 YISQYSEIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPKAVSKNNVIPALE---LVE 300
Db 154 -----MLDAIPNPLNKLKLVLPALPGAFAR 178
Qy 301 P-----IKKHEVPAKSDVYCEVEFLVKEVTKLIDNNKTEKILD-AFOKMCCKLPKSL 353
Db 179 PGPHQTDLSEQLPIPLP-FCWLKRTLKRVQAVI-----PKGVLAVALSVQVCHVPLV 232
Qy 354 SEECQEVVDYTGSIILLESPELVCSMLHCS-----GTRLPALTWHVTO-PKGG 407
Db 233 GGICQCLAERYTVLLDALLGRVVVQVGLVRLCRSTADAIQPALPALELIKWPLQDT 292

Qy 408 FCEVCKKLVGYLDRLNLEKSTKQEIILAALEKGC-SFLPDPYQKQCDQVAYEYEVLTIEL 466
Db 293 ECHFKCKSVI-----NQAWNTSEQAMPQAMHQAQLRFLWD--RQKCEQVQEHMPQLLALV 345
Qy 467 VEVNDSFVCLKIGACPSAHKPLLGTEKCIWGP 499
Db 346 PRSQDAHTSCQALGVCEAPASPL-----QCFQTP 374

RESULT 7
LNHUB
pulmonary surfactant protein B precursor [validated] - human
N:Alternate names: pulmonary surfactant proteolipid SP-B; pulmonary surfactant-associated protein B
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 08-Dec-2000
C:Accession: A31361; A28461; A27794; A27592; J00162; S21134
R:Pilot-Matias, T.J.; Kister, S.E.; Fox, J.L.; Kropp, K.; Glasser, S.W.; Whitsett, J.A.
DNA 8, 75-86, 1989
A>Title: Structure and organization of the gene encoding human pulmonary surfactant protein B
A:Reference number: A31361; MUID:89170128; PMID:2924687
A:Accession: A31361
A:Molecule type: DNA
A:Residues: 1-381 <PIL>
A:Cross-references: GB:M24461
A>Note: the codon given for residue 131 (ATT) is inconsistent with the authors' translation
R:Jacobs, K.A.; Phelps, D.S.; Steinbrink, R.; Fisch, J.; Kriz, R.; Mitsock, L.; Dougherty, J.
J. Biol. Chem. 262, 9808-9811, 1987
A>Title: Isolation of a cDNA clone encoding a high molecular weight precursor to a 6-kDa pulmonary surfactant protein B
A:Reference number: A28461; MUID:87250653; PMID:3597440
A:Accession: A28461
A:Molecule type: mRNA
A:Residues: 1-227, 'A', 229-381 <JAC>
A:Cross-references: GB:J02761; NID:G190673; PIDN:AAA60212.1; PID:G190674
A>Note: part of this sequence, including the amino end of the mature protein, was confirmed by R. Glasser, S.W.; Korfhaugen, T.R.; Weaver, T.; Pilot-Matias, T.; Fox, J.L.; Whitsett, J.A.
Proc. Natl. Acad. Sci. U.S.A. 84, 4007-4011, 1987
A>Title: cDNA and deduced amino acid sequence of human pulmonary surfactant-associated protein B
A:Reference number: A27794; MUID:87231940; PMID:3035561
A:Accession: A27794
A:Molecule type: mRNA
A:Residues: 'EFR', 99-317, 'L', 319-381 <GLA>
A:Cross-references: GB:M16764; NID:G338410; PIDN:AAA88099.1; PID:G338411
A>Note: 131-Ile was also found
A>Note: part of this sequence, including the amino end of the mature protein, was confirmed by R. Revak, S.D.; Merritt, T.A.; Degryse, E.; Stefani, L.; Courtney, M.; Hallman, M.; Cochran, J.
J. Clin. Invest. 81, 826-833, 1988
A>Title: Use of human surfactant low molecular weight apoproteins in the reconstitution of pulmonary surfactant
A:Reference number: A27592; MUID:88139786; PMID:3343343
A:Accession: A27592
A:Molecule type: mRNA
A:Residues: 139-177, 'V', 179-227, 'A', 228-381 <REV>
A:Cross-references: GB:M19097
A>Note: part of this sequence, including the amino end of the mature protein, was confirmed by R. Mizumoto, M.; Adachi, H.
Sapporo Igaku Zasshi 56, 731-742, 1987
A>Title: Primary structure of a hydrophobic 6kDa apoprotein (SP6) of human pulmonary surfactant
A:Reference number: J00162
A:Accession: J00162
A:Molecule type: protein
A:Residues: 201-207, 'X', 209-210, 'X', 212-227, 'A', 229-234, 'X', 236-245, 'X', 247, 'L', 249-253, 'R'; Johansson, J.; Joernvall, H.; Curstedt, T.
FEBS Lett. 301, 165-167, 1992
A>Title: Human surfactant polypeptide SP-B. Disulfide bridges, C-terminal end, and peptide
A:Reference number: S21134; MUID:92233937; PMID:1568474
A:Accession: S21134
A>Status: preliminary
A:Molecule type: protein
A:Residues: 201-227, 'I', 229-279 <JOH>
A>Note: 228-Ala was also found
C:Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lowers the

[illegible]

QY 244 YISQYSEIAIOMMMHQPKEICALVGFCEVDEKEMPMQTLVPAKASKNVIPALELVEPIK 303
Db 176 -----HTQ-----DLSAQPFPIPLP-----190
QY 304 KHEVPAKSDVYCEVCFVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSECEVVD 362
Db 191 -----LWLCRTLLKRIQAMI-----PKGVLAAMAVQVCHVPLVVGIGICQCLAE 235
QY 363 TYGSSILSILLEEVSPVCSMLHLCS-----GTRLPALTIVHTQ---PRDGGFCVCKKL 415
Db 236 RYTVILLEVLGLVLPQVGLVLRCSVSDSIGOVPTLEALPGEWLPQDPE-CHLQMSV 294
QY 416 VGYLDRLNLEKNSKQKQILAA-----LEKGCFLPDYKQCDQFVAEYEPVLIIELVEM 470
Db 295 TTQA-RNISEQTRPQAVHACLSQLDK-----OECEQFVAHAAP-----AA 335
QY 471 DPFVCLKIGACPSAHKPLLTGE-----KCIWGPSY 501
Db 336 EPAVOGL-----GCP-----RNLPGPGRVWATLSLQCIQSPHF 370

RESULT 13

T46069

hypothetical protein T18N14.110 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C:Accession: T46069

R:Delaney, M.; Berger, C.; Cooke, R.; Grellet, F.; Laudie, M.; Mewes, H.W.; Lemcke, K.;

submitted to the Protein Sequence Database, December 1999

A:Reference number: Z23013

A:Accession: T46069

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-213

A:Cross-references: EMBL:AL132968

A:Experimental source: cultivar Columbia; BAC clone T18N14

C:Genetics:

A:Map position: 3

A:Introns: 31/1; 80/3; 146/3; 166/3

A:Note: T18N14.110

Query Match 7.6%; Score 213; DB 2; Length 213;
Best Local Similarity 29.7%; Pred. No. 5.1e-07;
Matches 57; Conservative 35; Mismatches 90; Indels 10; Gaps 6;
QY 292 VIPALELVEPIKHEVPAKSDVYCEVCFVKEVTKLIDNNKTEKILDAFDKWCCKLPK 351
Db 18 VSDARSFVDSITSEKVNKEDV-CTLCCEVYTDALSYLEKKNVTQAEIIEIDLHRCSQL-R 75
QY 352 SLSECEQVVDYTGSSILSILLEEVSPVCSMLHLCSGTRLPALTIVHTQPKDGGFCV 411
Db 76 GYSQQCISLVDDY-VPLFFQLQSFPHYFCKRMNLG--KVALVEEARQDS----CGV 128
QY 412 CKKLGVYLDRLNLEKNSKQKQILAALEKGCFLPDYKQCDQFVAEYEPVLIIELVEM 471
Db 129 CHRTVSEILIKLPDPTDLDIVELLKGCCKSLKN-YEKKCTLVFEYGLIILVNAEEFLV 187
QY 472 PSFVCLKIGACP 483
Db 188 KNDVCTLLRACP 199

RESULT 14

T48201

hypothetical protein T20L15.70 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T48201

R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkee, W.; Stiekema, W.; Bancroft, I.; Mew

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24488

A:Accession: T48201

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-217 <BEV>
A:Cross-references: EMBL:AL162351
A:Experimental source: cultivar Columbia; BAC clone T20L15
C:Genetics:
A:Map position: 5
A:Introns: 30/1; 79/3; 146/3; 166/3
A:Note: T20L15.70

Query Match 7.3%; Score 203.5; DB 2; Length 217;
Best Local Similarity 27.8%; Pred. No. 2.2e-06;
Matches 52; Conservative 40; Mismatches 82; Indels 13; Gaps 7;
QY 298 LVEPIKK-HEVPAKSDVYCEVCFVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSE 356
Db 25 LLEPPESAHD-----DNQVCELCDDKYVTLVIDYLDQYDNQNELVEALHISCSQIP-PLKKQ 79
QY 357 COEVVDYTGSSILSILLEEVSPVCSMLHLCSGTRLPALTIVHTQPKDGGFCVCKKL 416
Db 80 CLSMVDHY-TQLFFTOVSTIKSDQICKRLNLQAV-TPAFASQVHQ-----GNCEACRETV 133
QY 417 GYLDRLNLEKNSKQKQILAALEKGCFLPDYKQCDQFVAEYEPVLIIELVEMDPSPFVC 476
Db 134 SEVTKLKPDPETKLKIIRLLLECKESL--NNYQDKCKGMVFEYGLMLTDLQKFLKKDVC 192
QY 477 LKIGACP 483
Db 193 TILHVCP 199

RESULT 15

T15677

hypothetical protein C28C12.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15677

R:Miller, N.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid C28C12.

A:Reference number: Z18387

A:Accession: T15677

A:Status: preliminary; translated from GB/EMBL/DDJ

A:Molecule type: DNA

A:Residues: 1-402 <MIL>

A:Cross-references: EMBL:U40797; NID:g1065916; PID:g1065921; PIDN:AAB37548.1; GSPDB:GN001

A:Experimental source: strain Bristol N2; clone C28C12

C:Genetics:

A:Gene: CESP:C28C12.5

A:Map position: 4

A:Introns: 29/3; 82/3; 124/3; 151/3; 258/3; 318/1; 372/3; 400/3

Query Match 6.3%; Score 176.5; DB 2; Length 402;
Best Local Similarity 20.3%; Pred. No. 0.00029;
Matches 75; Conservative 58; Mismatches 156; Indels 81; Gaps 15;
QY 184 GPRSKPQKNGDVCCDCIOMVTDIQTAVRTNFTFQALVEHVKEECDRLGPGMADICKN 243
Db 16 GAQSAPSP-----CESCKSNVQNFIDASKDRMMAQLKVS--LSMLC--VGTSHQSDCSK 66
QY 244 YISQYSEIAIOMMMHMQPKEICALVGFCEVDEKEMPMQTLVPAKASKNVIPALELVEPI 302
Db 67 TLDKLDPIAYKLAPYLADTSVAVCSKLMQCGESQFSLARLAMLKXSEIVA---NDNI 133
QY 303 KHEVPAKSDVYCEVCFVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSECEVVD 362
Db 124 MRQEV-----CDECOASTAQIGKLVGDEFTTYAVKTLQREVCVKSAGKAHKAC---N 172
QY 363 TYGSSILSILLEEV-----SPELVCSMLHLCSGTRLPALTIVHTQPKDGGFCVCKLV 417
Db 173 IFVSSVIPDLMTWKDMFTKELMCSNMGLCSATSKFAAREAPKQPAS-----EMWKSMM--- 226
QY 418 YLDRLNLEKNSKQKQILAALEKGCFLPDYKQCDQFVAEYEPVLIIELVEMDPSPFVC 450

```
Db      227  ---GMVKTSGBELMSCFE--CTLSADALLQEFIDKRQGTADDIQTVACNKMVANWTDG 280
Qy      451  CDOFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAHKPLLG---TEKCIWGPSYWCNTE 507
Db      281  CNDFVHMYMSTVLELTYNQDGRGICTTMMHSCCKENALVEMAMSEKVM/LG-----CEN-- 334
Qy      508  TAAQCNAVEH 517
Db      335  ---CKAVEH 340
```

Search completed: May 5, 2004, 13:34:33
Job time : 21.9003 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:14:53 ; Search time 55.7341 Seconds

(without alignments)
2606.077 Million cell updates/sec

Title: US-09-743-684a-1

Perfect score: 2789
Sequence: 1 MYALFLASLLGALAGPVL.....NTETRAQCNAYEHCKRHVN 524

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 27189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09C_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2789	100.0	524	9	US-09-870-759-60 Sequence 60, Appl
2	2789	100.0	524	10	US-09-751-708A-60 Sequence 36, App
3	2789	100.0	524	12	US-10-267-502-386 Sequence 61, Appl
4	2777.5	99.6	527	9	US-09-870-759-61 Sequence 61, Appl
5	2777.5	99.6	527	10	US-09-751-708A-61 Sequence 61, Appl
6	2777.5	99.6	527	14	US-10-060-036-73 Sequence 73, Appl
7	2757.5	98.9	523	9	US-09-767-007A-2 Sequence 40, Appl
8	2449.5	87.8	479	10	US-09-978-418-40 Sequence 17, App
9	1969	70.6	554	14	US-10-205-194-176 Sequence 1, Appl
10	1201	43.1	521	15	US-10-276-162-1 Sequence 90, App
11	1139.5	40.3	531	11	US-09-833-245-903 Sequence 340, App
12	1093	39.2	209	14	US-10-043-487-340 Sequence 8, Appl
13	834.5	29.9	362	12	US-10-332-426-8 Sequence 38, App
14	552	19.8	953	12	US-10-267-502-385 Sequence 4529, Ap
15	545	19.5	241	15	US-10-108-260A-4529

16	539.5	19.3	156	12	US-09-925-298-644 Sequence 644, App
17	539.5	19.3	156	14	US-10-102-806-644 Sequence 644, App
18	412	14.8	80	9	US-09-767-007A-3 Sequence 3, Appl
19	412	14.8	80	9	US-09-753-126-3 Sequence 3, Appl
20	412	14.8	80	15	US-10-330-697-3 Sequence 4, Appl
21	412	14.8	592	9	US-09-753-126-4 Sequence 10, Appl
22	412	14.8	592	15	US-10-330-697-4 Sequence 1223, Ap
23	349.5	12.5	381	15	US-10-236-031B-10 Sequence 276328, Ap
24	349.5	12.5	381	15	US-10-295-027-1223 Sequence 1, Appl
25	235.5	8.4	216	12	US-10-424-599-276328 Sequence 157904, A
26	211	7.6	40	10	US-09-780-438A-1 Sequence 41077, A
27	211	7.6	212	12	US-10-424-599-157904 Sequence 2, Appl
28	204.5	7.3	246	12	US-10-425-114-41077 Sequence 157906, A
29	201	7.2	38	10	US-09-780-438A-2 Sequence 5606, A
30	201	7.2	212	12	US-10-424-599-157906 Sequence 67929, A
31	197	7.1	252	12	US-10-425-114-55606 Sequence 51535, A
32	184	6.6	243	12	US-10-425-114-67929 Sequence 227975, A
33	175.5	6.3	265	12	US-10-425-114-51535 Sequence 12, Appl
34	169.5	6.1	236	12	US-10-424-599-227975 Sequence 60351, A
35	168.5	6.0	181	8	US-08-488-123-12 Sequence 3, Appl
36	165.5	5.9	246	12	US-10-425-114-60351 Sequence 1, Appl
37	163	5.8	514	14	US-10-339-351-3 Sequence 205602, A
38	162	5.8	514	14	US-10-339-351-1 Sequence 245585, A
39	159.5	5.7	281	12	US-10-424-599-205602 Sequence 71427, A
40	157.5	5.6	282	12	US-10-424-599-205603 Sequence 65142, A
41	156.5	5.6	507	12	US-10-424-599-245585 Sequence 40454, A
42	152	5.4	471	12	US-10-425-114-71427 Sequence 57134, A
43	151	5.4	393	12	US-10-425-114-65142 Sequence 57134, A
44	147.5	5.3	223	12	US-10-425-114-40454 Sequence 57134, A
45	146.5	5.3	295	12	US-10-425-114-57134

ALIGNMENTS

US-09-870-759-60
Sequence 60, Application US/09870759
Patent No. US20020177551A1
GENERAL INFORMATION:
APPLICANT: TERMAN, David S
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
FILE REFERENCE: 870759
CURRENT FILING DATE: 2002-01-14
PRIOR APPLICATION NUMBER: US 09/870,759
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 166
SOFTWARE: PatentIn version 3.1
SEQ ID NO 60
LENGTH: 524
TYPE: PRT
ORGANISM: Homo sapiens
US-09-870-759-60

Query Match 100.0%; Score 2789; DB 9; Length 524;
Best Local Similarity 100.0%; Pred. No. 3.6e-239;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MYALFLASLLGALAGPVLGLKECTRGSAWCVKTA	60	Sequence 60, Appl
DB	1	MYALFLASLLGALAGPVLGLKECTRGSAWCVKTA	60	Sequence 36, App
QY	61	LPDCICKDVTAAGDMLKDNATTEERILVLYLEKTC	120	Sequence 61, Appl
DB	61	LPDCICKDVTAAGDMLKDNATTEERILVLYLEKTC	120	Sequence 73, Appl
QY	121	IKGMSRGCEVSALNLCESLOKHAELNHOKLESNNKI	180	Sequence 1, Appl
DB	121	IKGMSRGCEVSALNLCESLOKHAELNHOKLESNNKI	180	Sequence 90, App
QY	181	PODGRSPROPKNDNDVCDPCICQMTDITQVTRTSTF	240	Sequence 340, App
DB	181	PODGRSPROPKNDNDVCDPCICQMTDITQVTRTSTF	240	Sequence 8, Appl

Db 181 PDGPRSKPQPKNDGVCDQDCIQMWTDIQTAVRNTSTFVQALVHVHKEECDRLGFGMADI 240
Qy 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCEVKEKMPQTLVPARVASKNVLPALVELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCEVKEKMPQTLVPARVASKNVLPALVELVE 300
Qy 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSECCQEV 360
Db 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSECCQEV 360
Qy 361 VDTYGSISILILEEVSPELVCSMLHLCSGTRLPALTVAHTVTPKDGFCCEVCKKLVGYLD 420
Db 361 VDTYGSISILILEEVSPELVCSMLHLCSGTRLPALTVAHTVTPKDGFCCEVCKKLVGYLD 420
Qy 421 RNLEKSTKOEILALEKSGSFLPDYOKQCDQFAVEYEPVLEILVEVMDPSFVCLKTIG 480
Db 421 RNLEKSTKOEILALEKSGSFLPDYOKQCDQFAVEYEPVLEILVEVMDPSFVCLKTIG 480
Qy 481 ACPSAHKPLLTGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLTGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524

RESULT 2

US-09-751-708A-60
; Sequence 60, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-751-708A-60

Query Match 100.0%; Score 2789; DB 10; Length 524;
Best Local Similarity 100.0%; Pred. No. 3,66-239;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYALFLASILGAALAGPVILGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKS 60
Db 1 MYALFLASILGAALAGPVILGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKS 60
Qy 61 LPDICKDVVTAAGDMLKDNATEEBEILVYLEKTCQDLPRKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPDICKDVVTAAGDMLKDNATEEBEILVYLEKTCQDLPRKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGEMSRPGEVCSALNLCESLQKHIAELNHQKQLESNNKIPELDMEVVAPEFMANIPILLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHIAELNHQKQLESNNKIPELDMEVVAPEFMANIPILLY 180
Qy 181 PDGPRSKPQPKNDGVCDQDCIQMWTDIQTAVRNTSTFVQALVHVHKEECDRLGFGMADI 240
Db 181 PDGPRSKPQPKNDGVCDQDCIQMWTDIQTAVRNTSTFVQALVHVHKEECDRLGFGMADI 240
Qy 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCEVKEKMPQTLVPARVASKNVLPALVELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCEVKEKMPQTLVPARVASKNVLPALVELVE 300
Qy 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSECCQEV 360
Db 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSECCQEV 360
Qy 361 VDTYGSISILILEEVSPELVCSMLHLCSGTRLPALTVAHTVTPKDGFCCEVCKKLVGYLD 420
Db 361 VDTYGSISILILEEVSPELVCSMLHLCSGTRLPALTVAHTVTPKDGFCCEVCKKLVGYLD 420

Db 361 VDTYGSISILILEEVSPELVCSMLHLCSGTRLPALTVAHTVTPKDGFCCEVCKKLVGYLD 420
Qy 421 RNLEKSTKOEILALEKSGSFLPDYOKQCDQFAVEYEPVLEILVEVMDPSFVCLKTIG 480
Db 421 RNLEKSTKOEILALEKSGSFLPDYOKQCDQFAVEYEPVLEILVEVMDPSFVCLKTIG 480
Qy 481 ACPSAHKPLLTGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLTGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524

RESULT 3

US-10-267-502-386
; Sequence 386, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: USD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 386
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-267-502-386

Query Match 100.0%; Score 2789; DB 12; Length 524;
Best Local Similarity 100.0%; Pred. No. 3,66-239;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALFLASILGAALAGPVILGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKS 60
Db 1 MYALFLASILGAALAGPVILGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKS 60
Qy 61 LPDICKDVVTAAGDMLKDNATEEBEILVYLEKTCQDLPRKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPDICKDVVTAAGDMLKDNATEEBEILVYLEKTCQDLPRKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGEMSRPGEVCSALNLCESLQKHIAELNHQKQLESNNKIPELDMEVVAPEFMANIPILLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHIAELNHQKQLESNNKIPELDMEVVAPEFMANIPILLY 180
Qy 181 PDGPRSKPQPKNDGVCDQDCIQMWTDIQTAVRNTSTFVQALVHVHKEECDRLGFGMADI 240
Db 181 PDGPRSKPQPKNDGVCDQDCIQMWTDIQTAVRNTSTFVQALVHVHKEECDRLGFGMADI 240
Qy 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCEVKEKMPQTLVPARVASKNVLPALVELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCEVKEKMPQTLVPARVASKNVLPALVELVE 300
Qy 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSECCQEV 360
Db 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSECCQEV 360
Qy 361 VDTYGSISILILEEVSPELVCSMLHLCSGTRLPALTVAHTVTPKDGFCCEVCKKLVGYLD 420
Db 361 VDTYGSISILILEEVSPELVCSMLHLCSGTRLPALTVAHTVTPKDGFCCEVCKKLVGYLD 420
Qy 421 RNLEKSTKOEILALEKSGSFLPDYOKQCDQFAVEYEPVLEILVEVMDPSFVCLKTIG 480
Db 421 RNLEKSTKOEILALEKSGSFLPDYOKQCDQFAVEYEPVLEILVEVMDPSFVCLKTIG 480
Qy 481 ACPSAHKPLLTGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLTGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524

RESULT 4
US-09-870-759-61


```
; Sequence 61, Application US/09870759
; Patent No. US2002017551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 61
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-870-759-61

Query Match      99.6%; Score 2777.5; DB 9; Length 527;
Best Local Similarity 99.4%; Pred. No. 3.9e-238;
Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MYALFLASLIGALAGVTLGKECTRGSAVWCQNVKTASDCGAVKHCLOQTWNKPTVKS 60
DB 1 MYALFLASLIGALAGVTLGKECTRGSAVWCQNVKTASDCGAVKHCLOQTWNKPTVKS 60
QY 61 LPDICKDVTTAAGDMLKDNATEBEILVYLEKTCMDLPKPMNSASCKEIVDSYLPVILI 120
DB 61 LPDICKDVTTAAGDMLKDNATEBEILVYLEKTCMDLPKPMNSASCKEIVDSYLPVILI 120
QY 121 IKGEMSRGVECSALNLCESLQKHLAELNHOKLESNKIPELDMEVVAPEPMANIPLLLY 180
DB 121 IKGEMSRGVECSALNLCESLQKHLAELNHOKLESNKIPELDMEVVAPEPMANIPLLLY 180
QY 181 PODGRSRKPOPKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVEHYKECDRLGPGMADI 240
DB 181 PODGRSRKPOPKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVEHYKECDRLGPGMADI 240
QY 241 CKNYISQYSEIATIQMMHMM--OPKEICALVGFCDVEKEMPMOTLVPAKVASKNVIIPALE 297
DB 241 CKNYISQYSEIATIQMMHMM--OPKEICALVGFCDVEKEMPMOTLVPAKVASKNVIIPALE 297
QY 298 LVEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKNCKSLPKSLSEEC 357
DB 298 LVEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKNCKSLPKSLSEEC 357
QY 301 LVEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKNCKSLPKSLSEEC 360
DB 301 LVEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKNCKSLPKSLSEEC 360
QY 358 QEVVDYTGSSILSLILEVSPELVCSMLHLCSGTRLPALTVHTQPKDGFCEVCKLVG 417
DB 358 QEVVDYTGSSILSLILEVSPELVCSMLHLCSGTRLPALTVHTQPKDGFCEVCKLVG 417
QY 418 YIDRLNLEKNSKTOEILALEKGCSPFLPDYKQCDQFVAEYEPVILIEILVEVMDPSFVCL 477
DB 418 YIDRLNLEKNSKTOEILALEKGCSPFLPDYKQCDQFVAEYEPVILIEILVEVMDPSFVCL 477
QY 421 YIDRLNLEKNSKTOEILALEKGCSPFLPDYKQCDQFVAEYEPVILIEILVEVMDPSFVCL 480
DB 421 YIDRLNLEKNSKTOEILALEKGCSPFLPDYKQCDQFVAEYEPVILIEILVEVMDPSFVCL 480
QY 478 KIGACPSAHKPLGTGTEKCIWGPSTWCONTEPAACNAVEHCKRHAWN 524
DB 478 KIGACPSAHKPLGTGTEKCIWGPSTWCONTEPAACNAVEHCKRHAWN 524
; RESULT 5
; US-09-751-708A-61
; Sequence 61, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: Patent in version 3.1
```

```
; SEQ ID NO 61
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-751-708A-61

Query Match      99.6%; Score 2777.5; DB 10; Length 527;
Best Local Similarity 99.4%; Pred. No. 3.9e-238;
Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MYALFLASLIGALAGVTLGKECTRGSAVWCQNVKTASDCGAVKHCLOQTWNKPTVKS 60
DB 1 MYALFLASLIGALAGVTLGKECTRGSAVWCQNVKTASDCGAVKHCLOQTWNKPTVKS 60
QY 61 LPDICKDVTTAAGDMLKDNATEBEILVYLEKTCMDLPKPMNSASCKEIVDSYLPVILI 120
DB 61 LPDICKDVTTAAGDMLKDNATEBEILVYLEKTCMDLPKPMNSASCKEIVDSYLPVILI 120
QY 121 IKGEMSRGVECSALNLCESLQKHLAELNHOKLESNKIPELDMEVVAPEPMANIPLLLY 180
DB 121 IKGEMSRGVECSALNLCESLQKHLAELNHOKLESNKIPELDMEVVAPEPMANIPLLLY 180
QY 181 PODGRSRKPOPKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVEHYKECDRLGPGMADI 240
DB 181 PODGRSRKPOPKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVEHYKECDRLGPGMADI 240
QY 241 CKNYISQYSEIATIQMMHMM--OPKEICALVGFCDVEKEMPMOTLVPAKVASKNVIIPALE 297
DB 241 CKNYISQYSEIATIQMMHMM--OPKEICALVGFCDVEKEMPMOTLVPAKVASKNVIIPALE 297
QY 298 LVEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKNCKSLPKSLSEEC 357
DB 298 LVEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKNCKSLPKSLSEEC 357
QY 301 LVEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKNCKSLPKSLSEEC 360
DB 301 LVEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFKNCKSLPKSLSEEC 360
QY 358 QEVVDYTGSSILSLILEVSPELVCSMLHLCSGTRLPALTVHTQPKDGFCEVCKLVG 417
DB 358 QEVVDYTGSSILSLILEVSPELVCSMLHLCSGTRLPALTVHTQPKDGFCEVCKLVG 417
QY 418 YIDRLNLEKNSKTOEILALEKGCSPFLPDYKQCDQFVAEYEPVILIEILVEVMDPSFVCL 477
DB 418 YIDRLNLEKNSKTOEILALEKGCSPFLPDYKQCDQFVAEYEPVILIEILVEVMDPSFVCL 477
QY 421 YIDRLNLEKNSKTOEILALEKGCSPFLPDYKQCDQFVAEYEPVILIEILVEVMDPSFVCL 480
DB 421 YIDRLNLEKNSKTOEILALEKGCSPFLPDYKQCDQFVAEYEPVILIEILVEVMDPSFVCL 480
QY 478 KIGACPSAHKPLGTGTEKCIWGPSTWCONTEPAACNAVEHCKRHAWN 524
DB 478 KIGACPSAHKPLGTGTEKCIWGPSTWCONTEPAACNAVEHCKRHAWN 524
; RESULT 6
; US-10-060-036-73
; Sequence 73, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-060-036-73

Query Match      99.6%; Score 2777.5; DB 14; Length 527;
Best Local Similarity 99.4%; Pred. No. 3.9e-238;
```

Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MYALFLASLLGALAGVTLGKECTRGSAYWCQNVKTASDCGAVKHCLQTVMNKPVTYS 60
 Db 1 MYALFLASLLGALAGVTLGKECTRGSAYWCQNVKTASDCGAVKHCLQTVMNKPVTYS 60

QY 61 LPDCICKDVVTAAGMDLKDNTATEEILVLEKTCMDLPRKPMNSASCKEIVDSYLPVILI 120
 Db 61 LPDCICKDVVTAAGMDLKDNTATEEILVLEKTCMDLPRKPMNSASCKEIVDSYLPVILI 120

QY 121 IKGEMSRGGEVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVAPPMANIPILLY 180
 Db 121 IKGEMSRGGEVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVAPPMANIPILLY 180

QY 121 IKGEMSRGGEVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVAPPMANIPILLY 180
 Db 121 IKGEMSRGGEVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVAPPMANIPILLY 180

QY 181 POGPRSKPQPKNDGVCODCIQWVTDIQTAVRTNSTFVQALVHVEKBCDRLGPMADI 240
 Db 181 POGPRSKPQPKNDGVCODCIQWVTDIQTAVRTNSTFVQALVHVEKBCDRLGPMADI 240

QY 241 CKNYISQYSEIAIOMMMHMOQPKELCALVGFCDVEKEMPMQTLVPAKASKNVPALE 297
 Db 241 CKNYISQYSEIAIOMMMHMOQPKELCALVGFCDVEKEMPMQTLVPAKASKNVPALE 297

QY 298 LVSEIKKHEVPASDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFPMKSKLPKSLSEEC 357
 Db 301 LVSEIKKHEVPASDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFPMKSKLPKSLSEEC 357

QY 358 QEVVDYGSLLSTLLEEVSPELVCSMLHCSGTRLPALTVAHTOPKDGFCVCKKLVG 417
 Db 361 QEVVDYGSLLSTLLEEVSPELVCSMLHCSGTRLPALTVAHTOPKDGFCVCKKLVG 417

QY 418 YIDRNEKSTKQEIILAEKGCFLPDPYQKOCDFVAEYEPVLEILVEMDPSFVCL 477
 Db 421 YIDRNEKSTKQEIILAEKGCFLPDPYQKOCDFVAEYEPVLEILVEMDPSFVCL 477

QY 478 KIGACPSAKPPLGTREKCIWPSYWCQNTETAQCAVAHCKRHAWN 524
 Db 481 KIGACPSAKPPLGTREKCIWPSYWCQNTETAQCAVAHCKRHAWN 524

RESULT 7
 US-09-767-007A-2
 ; Sequence 2, Application US/09767007A
 ; Patent No. US20020072727SAI
 ; GENERAL INFORMATION:
 ; APPLICANT: John S. O'Brien
 ; APPLICANT: Yasuo Kishimoto
 ; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
 ; FILE REFERENCE: MYELOS 2DC1C1
 ; CURRENT APPLICATION NUMBER: US/09/767,007A
 ; PRIOR FILING DATE: 2001-01-22
 ; PRIOR APPLICATION NUMBER: 08/958,970
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 523
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 ; US-09-767-007A-2

Query Match 98.9%; Score 2757.5; DB 9; Length 523;
 Best Local Similarity 99.2%; Pred. No. 2.3e-236;
 Matches 520; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 MYALFLASLLGALAGVTLGKECTRGSAYWCQNVKTASDCGAVKHCLQTVMNKPVTYS 60
 Db 1 MYALFLASLLGALAGVTLGKECTRGSAYWCQNVKTASDCGAVKHCLQTVMNKPVTYS 60

QY 61 LPDCICKDVVTAAGMDLKDNTATEEILVLEKTCMDLPRKPMNSASCKEIVDSYLPVILI 120
 Db 61 LPDCICKDVVTAAGMDLKDNTATEEILVLEKTCMDLPRKPMNSASCKEIVDSYLPVILI 120

QY 121 IKGEMSRGGEVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVAPPMANIPILLY 180
 Db 121 IKGEMSRGGEVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVAPPMANIPILLY 180

QY 181 POGPRSKPQPKNDGVCODCIQWVTDIQTAVRTNSTFVQALVHVEKBCDRLGPMADI 240
 Db 181 POGPRSKPQPKNDGVCODCIQWVTDIQTAVRTNSTFVQALVHVEKBCDRLGPMADI 240

QY 241 CKNYISQYSEIAIOMMMHMOQPKELCALVGFCDVEKEMPMQTLVPAKASKNVPALE 297
 Db 240 CKNYISQYSEIAIOMMMHMOQPKELCALVGFCDVEKEMPMQTLVPAKASKNVPALE 297

QY 301 PIKHEVPASDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFPMKSKLPKSLSEECOE 360
 Db 300 PIKHEVPASDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFPMKSKLPKSLSEECOE 359

QY 361 VDTYGSLLSTLLEEVSPELVCSMLHCSGTRLPALTVAHTOPKDGFCVCKKLVG 417
 Db 360 VDTYGSLLSTLLEEVSPELVCSMLHCSGTRLPALTVAHTOPKDGFCVCKKLVG 417

QY 421 RNEKSTKQEIILAEKGCFLPDPYQKOCDFVAEYEPVLEILVEMDPSFVCL 480
 Db 420 RNEKSTKQEIILAEKGCFLPDPYQKOCDFVAEYEPVLEILVEMDPSFVCL 479

QY 481 ACPSAHKPLGTREKCIWPSYWCQNTETAQCAVAHCKRHAWN 524
 Db 480 ACPSAHKPLGTREKCIWPSYWCQNTETAQCAVAHCKRHAWN 523

RESULT 8
 US-09-978-418-40
 ; Sequence 40, Application US/09978418
 ; Publication No. US20030118997A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Benjamin, Stephan
 ; APPLICANT: Tanaka, Hiroaki
 ; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
 ; FILE REFERENCE: 142 US REG
 ; CURRENT APPLICATION NUMBER: US/09/978,418
 ; CURRENT FILING DATE: 2002-10-15
 ; PRIOR APPLICATION NUMBER: 60/311,305
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: 60/314,734
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: 60/318,204
 ; PRIOR FILING DATE: 2001-09-07
 ; PRIOR APPLICATION NUMBER: 60/326,470
 ; PRIOR FILING DATE: 2001-10-01
 ; NUMBER OF SEQ ID NOS: 52
 ; SOFTWARE: JPatent
 ; SEQ ID NO 40
 ; LENGTH: 479
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-978-418-40

Query Match 87.8%; Score 2449.5; DB 10; Length 479;
 Best Local Similarity 99.4%; Pred. No. 5.1e-209;
 Matches 469; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MYALFLASLLGALAGVTLGKECTRGSAYWCQNVKTASDCGAVKHCLQTVMNKPVTYS 60
 Db 1 MYALFLASLLGALAGVTLGKECTRGSAYWCQNVKTASDCGAVKHCLQTVMNKPVTYS 60

QY 61 LPDCICKDVVTAAGMDLKDNTATEEILVLEKTCMDLPRKPMNSASCKEIVDSYLPVILI 120
 Db 61 LPDCICKDVVTAAGMDLKDNTATEEILVLEKTCMDLPRKPMNSASCKEIVDSYLPVILI 120

QY 121 IKGEMSRGGEVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVAPPMANIPILLY 180
 Db 121 IKGEMSRGGEVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVAPPMANIPILLY 180

QY 181 POGPRSKPQPKNDGVCODCIQWVTDIQTAVRTNSTFVQALVHVEKBCDRLGPMADI 240

```
Db 181 P Q D G R S K P Q P K D N G D V C D C I Q W T D I Q T A R T N S T F V Q L V E H V E K B C B L G P G M A D I 240
Qy 241 C K N Y I S Q Y S E I A I Q M M M H M --- O P R E I C A L V G F C D E V K E M P Q T L V P A K V A S K N I P A L E 297
Db 241 C K N Y I S Q Y S E I A I Q M M M H M D Q O P K E I C A L V G F C D E V K E M P Q T L V P A K V A S K N I P A L E 300
Qy 248 L V E P I K K H E V P A K S V Y C E V C E F L V K E V T K L I D N N K T E K E I I D A P D K M C S K L P K S L S E C 357
Db 301 L V E P I K K H E V P A K S V Y C E V C E F L V K E V T K L I D N N K T E K E I I D A P D K M C S K L P K S L S E C 360
Qy 358 Q E V V D T Y S S I L I L E E V S P E L V C S M L H C S G T R L P A L T V H T O P K D G F C E V C K L V G 417
Db 361 Q E V V D T Y S S I L I L E E V S P E L V C S M L H C S G T R L P A L T V H T O P K D G F C E V C K L V G 420
Qy 418 Y L D R L E K N S T K O E I L A L E K G C S F L P D P Y Q K C Q O F A E Y E P V L I E I L V E Y 469
Db 421 Y L D R L E K N S T K O E I L A L E K G C S F L P D P Y Q K C Q O F A E Y E P V L I E I L V E Y 472
```

RESULT 9
US-10-205-194-176
Sequence 176, Application US/10205194
Publication No. US20030134301A1
GENERAL INFORMATION:
APPLICANT: Warner-Lambert Company
APPLICANT: Lee, Kevin
APPLICANT: Dixon, Alistair
APPLICANT: Brooksbank, Robert
APPLICANT: Pincock, Robert
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
FILE REFERENCE: WL-A-018201
CURRENT APPLICATION NUMBER: US/10/205,194
CURRENT FILING DATE: 5200-07-24
PRIOR APPLICATION NUMBER: GB 0118354.0
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 177
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 176
LENGTH: 554
TYPE: PRT
ORGANISM: Rattus sp.
FEATURE:
OTHER INFORMATION: Prosaposin
US-10-205-194-176

```
Query Match 70.6%; Score 1969; DB 14; Length 554;  
Best Local Similarity 65.8%; Pred. No. 3,6e-166;  
Matches 365; Conservative 77; Mismatches 81; Indels 32; Gaps 3;  
Qy 1 M Y A L F L A S L G A L A G P L L G K E C T R G S A V C Q V N K T A S D C G A V K C L Q T W M K P T Y K S 60  
Db 1 M Y A L F L A S L V T A L T S P V O D P K I C S G S A V C R D V K T A V D C R A V H C Q Q W M S K P T A K S 60  
Qy 61 L P C D I C K O V V T A G M L K D N A T E E R I L V Y L E K T C D M L P R M N N S A S C K E I V D S Y L V I I D I 120  
Db 61 L P C D I C K V V T A G M L K D N A T E E R I L A Y L E K T C A M I H D S S A S C K E V D S Y L V I I D M 120  
Qy 121 I K G E M S R P E V C S A L N T C E S L O K H L A E L N H O K E L S N K I P E L D M T E V A P F M A N I P L L Y 180  
Db 121 I K G E M S N P E V C S A L N T C O S I Q E Y L A E Q N - Q R Q L E S N K I P E V D L A R V V A P F M S N I P L L Y 179  
Qy 181 P O D G R S K P Q P K D N G D V C D C I Q W T D I Q T A R T N S T F V Q L V E H V E K B C B L G P G M A D I 240  
Db 180 P O D R P R S O P Q P K A N D E V C O D C M K L V T D I Q T A V R T N S S F V Q G L V D H V K E D C D R L G G V S D I 239  
Qy 241 C K N Y I S Q Y S E I A I Q M M M H M O P K E I C A L V G F C D E V K E M P Q T L V P A K V A S K N I P A L E Y E 300  
Db 240 C K N Y V D Q I S E V A V Q M M M H M O P K E I C V M V G F C D E V K R V P K R I L V P A T E A I K N I L P A L E L T D 299  
Qy 301 P I K K H E V P A K S V Y C E V C E F L V K E V T K L I D N N K T E K E I I D A P D K M C S K L P K S L S E C D E V 360  
Db 300 P Y E Q V I Q A Q N V I F C Q V C Q L V M R K S E L I N N A T E E L I I K G L S K A C S L I P A P A S T K C Q E V 359
```

```
Qy 361 V D T Y S S I L I L E E V S P E L V C S M L H C S G ----- T R L P ----- 394  
Db 360 L V T F P S L I D V L M H E V N E N F I C G V I S L C S A N P N I V G T I E O P A A I V S A L P E R P A P K O P E 419  
Qy 395 ----- A L T V H T O P K D G F C E V C K K L V G Y L D R L E K N S T K O E I L A L E K G C S F L P D P Y Q K 449  
Db 420 E P K S A L A R A H P P Q K N G S F C E V C K K L V I Y L E H N E K N S T K E A V L A A L E K G C S F L P D P Y Q K 479  
Qy 450 Q C D O F V A E Y E P V L I E I L V E W D P S F V C L K I A C S A H K P L L G T E K T I G S Y W C O N T E T A 509  
Db 480 Q C D E F V A E Y E P L L E I L V E W D P S F V C S K I G V C S A Y L L G T E K C V A P G Y W C O N M E T A 539  
Qy 510 A O C N A V E H C R R H V M N 524  
Db 540 A R C N A V D H C K R H V M N 554
```

RESULT 10
US-10-276-162-1
Sequence 1, Application US/10276162
Publication No. US20030215822A1
GENERAL INFORMATION:
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: YAO, Monique G.
APPLICANT: BRUNS, Christopher M.
APPLICANT: YUE, Henry
APPLICANT: DELBEANE, Angelo M.
APPLICANT: HAFALIA, April
APPLICANT: PATTERSON, Chandra
APPLICANT: POLICKY, Jennifer L.
APPLICANT: TRIBOULEY, Catherine M.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: NGUYEN, Danielle B.
APPLICANT: LAU, Preci
APPLICANT: TANG, Y. Tom
APPLICANT: JACKSON, Jennifer L.
APPLICANT: LU, Dyung Aina M.
APPLICANT: BATRA, Sajeev
APPLICANT: AU-YOUNG, Janice
APPLICANT: REDDY, Roopa
APPLICANT: AZIMZAI, Yalda
TITLE OF INVENTION: SECRETED PROTEINS
FILE REFERENCE: PI-0071 USN
CURRENT APPLICATION NUMBER: US/10/276,162
CURRENT FILING DATE: 2002-10-15
PRIOR APPLICATION NUMBER: US01/11861
PRIOR FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 60/197,854
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/202,373
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/205,899
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: 60/210,155
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: 60/209,401
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PERL Program
SEQ ID NO: 1
LENGTH: 521
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Inocyte ID No. US20030215822A1 7473577CD1
US-10-276-162-1

Query Match 43.1%; Score 1201; DB 15; Length 521;
Best Local Similarity 44.3%; Pred. No. 7.1e-98;
Matches 237; Conservative 98; Mismatches 158; Indels 42; Gaps 10;

```

Qy 3 ALFLASLLGALAGPVLGLKECTRGSAVWCQNVKTAASDCGAVKGLQTVWKPETYKSLP 62
Db 4 ALILLPBLGATTAASPTSGPOEAKGSGTWQCQDLQTAARCAVGYCOGAWMNPPTAKSLP 63
Qy 63 CDICKDVVTAAGDMLKDNATEEBELVYLEKTCMDLPKPNMSASCKEIVDSYLPVILDIK 122
Db 64 CDVQDIAAAGANGLNPDATESDILALVMKTCMLPSQESSAGCKMNVDAHSAILSMIR 123
Qy 123 GEM-SRPEVCASALNLCESLOKHLAEINHOKESENKIPBLDMTEVAVPAPMANIPLLLYP 181
Db 124 GAPDSAPAOVCTALSLCEPLQRLHATL-----RLSKEDTTEAVAFPMANGPLTFHP 175
Qy 182 QDPRSKQPKDNDGVQDCIQMTVDIQTAVRTNSFTVQALVHVKEECDRLGPMADIC 241
Db 176 RQAP-----EGALQDCQVVRQVSRLOEAVRSNLTLDL---NIQOCESLGFGLAVLC 224
Qy 242 KNTISQYSEIALIOMMMHMQPKICALVGFCEVKEKEMQTLVPAK---VASKNVIPALEL 298
Db 225 KNTLFQFVPADQALRLPLPOELCRKGFCEELG-----APARLTQVAMGVBSLEL 277
Qy 299 VEPKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSKLPKSLSECO 358
Db 278 GLPRKQSEMOKAGVTCVCANVQKLDHMLNSSELMTTHALERVCSVMPASITKECI 337
Qy 359 EVDVYGSILSLILEEVSPELVCSMLHLCSGTRLPALTVE-----VTQPK-----DGGF 408
Db 338 ILVDYSPSLVQ-LVAKITPEKVKCFIRLC-GNRRRAVADAVAIVSPEDMAENQSGF 395
Qy 409 CEVCKKLVGYLDNRLEKNSTKOEILALEKSGSFLDPYQKQCDQFVAEYEPVLEILVE 468
Db 396 CNGCKRLTVSSHNSKSTKRDILVAFKGGCSILPLPYMIQCKHFVVOYEPVLESLKD 455
Qy 469 VMDPSFVCLIKIGACPSAHKPLLTGTEKCIWGPSTWCONTEAQAQNAVEHCKRHVW 523
Db 456 WMDPVAVCKVGAHGPRTPPLIGTDQCALGSPFWCRSGQEAALCNVAHGCQKHVW 510

```

RESULT 11

US-09-833-245-903

Sequence 903, Application US/09833245

Publication No. US20040010134A1

GENERAL INFORMATION:

APPLICANT: Human Genome Sciences, Inc.

TITLE OF INVENTION: Alkylmucin Fusion Proteins

FILE REFERENCE: PFS46PCT

CURRENT APPLICATION NUMBER: US/09/833,245

CURRENT FILING DATE: 2001-04-12

PRIOR APPLICATION NUMBER: 60/229, 358

PRIOR FILING DATE: 2000-04-12

PRIOR APPLICATION NUMBER: 60/256, 931

PRIOR FILING DATE: 2000-12-21

PRIOR APPLICATION NUMBER: 60/199, 384

PRIOR FILING DATE: 2000-04-25

NUMBER OF SEQ ID NOS: 2267

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 903

LENGTH: 531

TYPE: PRT

ORGANISM: Homo sapiens

US-09-833-245-903

```

Query Match 40.9%; Score 1139.5; DB 11; Length 531;
Best Local Similarity 43.2%; Pred. No. 2.2e-92;
Matches 231; Conservative 99; Mismatches 162; Indels 43; Gaps 11;
Qy 3 ALFLASLLGALAGPVLGLKECTRGSAVWCQNVKTAASDCGAVKGLQTVWKPETYKSLP 62
Db 4 ALILLPBLGATTAASPTSGPOEAKGSGTWQCQDLQTAARCAVGYCOGAWMNPPTAKSLP 63
Qy 63 CDICKDVVTAAGDMLKDNATEEBELVYLEKTCMDLPKPNMSASCKEIVDSYLPVILDIK 122
Db 64 CDVQDIAAAGANGLNPDATESDILALVMKTCMLPSQESSAGCKMNVDAHSAILSMIR 123

```

```

Qy 123 GEM-SRPEVCASALNLCESLOKHLAEINHOKESENKIPBLDMTEVAVPAPMANIPLLLYP 181
Db 124 GAPDSAPAOVCTALSLCEPLQRLHATL-----RLSKEDTTEAVAFPMANGPLTFHP 175
Qy 182 QDPRSKQPKDNDGVQDCIQMTVDIQTAVRTNSFTVQALVHVKEECDRLGPMADIC 241
Db 176 RQAP-----EGALQDCQVVRQVSRLOEAVRSNLTLDL---NIQOCESLGFGLAVLC 224
Qy 242 KNTISQYSEIALIOMMMHMQPKICALVGFCEVKEKEMQTLVPAK---VASKNVIPALEL 298
Db 225 KNTLFQFVPADQALRLPLPOELCRKGFCEELG-----APARLTQVAMGVBSLEL 277
Qy 299 VEPKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSKLPKSLSECO 358
Db 278 GLPRKQSEMOKAGVTCVCANVQKLDHMLNSSELMTTHALERVCSVMPASITKECI 337
Qy 359 EVDVYGSILSLILEEVSPELVCSMLHLCSGTRLPALTVE-----VTQPK-----DGGF 408
Db 338 ILVDYSPSLVQ-LVAKITPEKVKCFIRLC-GNRRRAVADAVAIVSPEDMAENQSGF 395
Qy 409 CEVCKKLVGYLDNRLEKNSTKOEILALEKSGSFLDPYQKQCDQFVAEYEPVLEILVE 468
Db 396 CNGCKRLTVSSHNSKSTKRDILVAFKGGCSILPLPYMIQCKHFVVOYEPVLESLKD 455
Qy 469 VMDPSFVCLIKIGACPSAHKPLLTGTEKCIWGPSTWCONTEAQAQNAVEHCKRHVW 523
Db 456 WMDPVAVCKVGAHGPRTPPLIGTDQCALGSPFWCRSGQEAALCNVAHGCQKHVW 509

```

RESULT 12

US-10-043-487-340

Sequence 340, Application US/10043487

Publication No. US20030055220A1

GENERAL INFORMATION:

APPLICANT: HYBRIGENICS

TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypeptide

FILE REFERENCE: B4778A

CURRENT APPLICATION NUMBER: US/10/043,487

CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/261,130

PRIOR FILING DATE: 2001-01-12

NUMBER OF SEQ ID NOS: 561

SOFTWARE: PatentIn version 3.1

SEQ ID NO 340

LENGTH: 209

TYPE: PRT

ORGANISM: Shigella flexneri

US-10-043-487-340

```

Query Match 39.2%; Score 1093; DB 14; Length 209;
Best Local Similarity 100.0%; Pred. No. 7.5e-89;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 151 OKQLESNKIPBLDMTEVAVPAPMANIPLLLYPODGRSKROPKNDGVQDCIQMTVDIQT 210
Db 1 OKQLESNKIPBLDMTEVAVPAPMANIPLLLYPODGRSKROPKNDGVQDCIQMTVDIQT 60
Qy 211 AVRTNSTFVQALVHVKEECDRLGPMADICKNTISQYSEIALIOMMMHMQPKICALVGF 270
Db 61 AVRTNSTFVQALVHVKEECDRLGPMADICKNTISQYSEIALIOMMMHMQPKICALVGF 120
Qy 271 CDEVKEMQTLVPAKVASKNVIPALELVEPIKKEHVPKSDVYCEVCEFLVKEVTKLID 330
Db 121 CDEVKEMQTLVPAKVASKNVIPALELVEPIKKEHVPKSDVYCEVCEFLVKEVTKLID 180
Qy 331 NNKTEKEILDAFDKMSKLPKSLSECOE 359
Db 181 NNKTEKEILDAFDKMSKLPKSLSECOE 209

```

RESULT 13

Query	Database	Score	Length	Gap
412	CKKLVGYLDNRNLEKXSTKQEIILAALEKCSFLPRPYQKQCDQFAAEYEPVLIIEIVMD	19.8%	552	DB 12
240	CKKLLTVSSHHNLESKSTKRDILIAVKGCSSTLPRLPYMIQCHGFYQYEPVLIISLKQMD	23.1%	552	DB 12
472	PSFVCLIKAGCSAHPFLLTGTEKCIWGSYVCONTEPAQCAVAHEKCRHW	23.1%	552	DB 12
300	PVAVCKKVGACGPRFTPLLTGTCALGFSFMCRSQEAACLNVAHQKQKHW	23.1%	552	DB 12
US-10-267-502-385	Sequence 385, Application US/10267502			
US-10-267-502-385	Publication No. US20040071700A1			
GENERAL INFORMATION:				
APPLICANT:	Kim, Jaeseob			
TITLE OF INVENTION:	Gallant, Ron			
FILE REFERENCE:	LSD-07416			
CURRENT FILING DATE:	2003-01-27			
NUMBER OF SEQ ID NOS:	439			
SOFTWARE:	Patentin version 3.2			
SEQ ID NO 385				
LENGTH:	953			
TYPE:	PRT			
ORGANISM:	Drosophila melanogaster			
US-10-267-502-385				
Query Match	19.8% Score 552; DB 12; Length 953;			
Best Local Similarity	23.1%; Pred. No. 9.6e-40;			
Matches	147; Conservative 107; Mismatches 220; Indels 162; Gaps 199			
1	MYALPFLASLGA-ALAGFVLGLKECTGSAVWCQNVYASDCGAVKGLQTWNK---P	56		
6	LIALVALCAFGVPAAPATPLTSSKCTGSPYWGQNFNSKECATHTCIGTWETQKVP	65		
57	TVKSLPDCIDKDVYTAGMCLKONATEEELVYLEKTCDWLPKFNMSAKCEIYDSTLPV	116		
66	VDTDSICTTCKMVTQARDQLKSNQTEBELKEVEFGSKLPIPIKIQEKCIKAVDPLPE	125		
117	ILDIINGEMSRPEVQSAALNLCES-----LQKH-----	144		
126	LVEALSAQNM-PRQVCSVAGLCNSARIDELYKNGIQAGLDTVONEDDSSEETELAMQN	184		
145	-----LAEINHQKQLSESKNIPELDMTEV-----APFNANIPLLP-----	179		
185	QLSCGNONLSTRMMSKFAATDRD---DMVETMLHMGSSLSFSFDACNIVLYTFNDIYD	241		
180	-----YPQDGPSSKQPK---DNCD-----VCQDCIQMTDIQ	209		
242	HVSKHLTTDAVCHVSGVCSARYHQHEEKKQPEALVALDADDDIPCEILCEQLVGNLIDL	301		
210	TAVRTNSTVQALVHEVKECDRLGPMADIQKAVISQYSIAIQ-NMMHQPKELCALV	268		
302	VANTTETTERKQWEGFCQD---SKGRKDCLSITVDYVIVVITYETLVSKIDANGACMI	356		
269	GFCDL-----VEMPEQTLVPAKVASKVIVALEL-----VEEIKKGE-----	306		
357	GICQKNSASMSKDVPIMLLP-----VIERAQVITIEKLEHGEKKQGLGASPRKSQ	409		
307	-----VP-----AKSDVCEVCEFLVKEVTKLIDNNKTEKELIDAFMCS	347		
410	EIIDMQLPTIDHMGANPGALVEGGELCTLCEYMLHFIQETLAPSTPDEIKHTVENICA	469		
348	KLPKSLSECGQVNDPYGSSILSTILLEVSPVLSYMSLHICSGRLRALYVHTQP-----	403		
470	KLPSSVAGGCRNFVEVYGDVATLALVQGLNPRDVCPLMQMKPKLPEKEDVEVFNQPAS	529		
404	--KDGFCFCVCKKLVGYLDNRNLEKXSTKQEIILAALEKCSFLPRPYQKQCDQFAAEYEPV	461		
530	DEDDPPTFCLEPAVEQAOQMKIRDMKSKDNIKKYLVNGICSLHPRNIEECVDVFNYYISNB	589		
462	LIEILVEMDPSPFVCLIKAGCSAHPFLLTGTEKCIW	97		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:10:37 ; Search time 22.8908 Seconds
(without alignments)
1181.786 Million cell updates/sec

Title: US-09-743-684A-1

Perfect score: 2789
Sequence: 1 MYAFLPLASLIGALAGPVL.....NTETRAQCNAYEHCKRHWN 524

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2789	100.0	524	4	US-09-352-548-1
2	2766.5	99.2	523	1	US-08-100-247-2
3	2766.5	99.2	523	4	US-08-756-031-2
4	2759.5	98.9	523	1	US-08-232-513A-3
5	2757.5	98.9	523	1	US-08-483-146A-2
6	2757.5	98.9	523	1	US-08-484-594A-2
7	2757.5	98.9	523	4	US-09-076-258A-2
8	431	15.5	81	4	US-09-352-548-2
9	418.5	15.0	80	2	US-08-584-671-15
10	418.5	15.0	80	3	US-09-027-376-15
11	415.5	14.9	80	3	US-09-094-192-15
12	412	14.8	80	1	US-08-100-247-3
13	412	14.8	80	1	US-08-483-146A-3
14	412	14.8	80	1	US-08-232-513A-4
15	412	14.8	80	1	US-08-484-594A-3
16	412	14.8	80	4	US-09-076-258A-3
17	412	14.8	80	4	US-08-756-031-3
18	350.5	12.6	381	2	US-09-193-877-2
19	327	11.7	61	2	US-08-584-671-13
20	327	11.7	61	3	US-09-027-376-13
21	327	11.7	61	3	US-09-094-192-13
22	327	11.7	69	3	US-09-268-070-2
23	322	11.5	79	2	US-08-584-671-16
24	322	11.5	79	3	US-09-027-376-16
25	322	11.5	79	3	US-09-094-192-16
26	321	11.5	79	2	US-08-584-671-14
27	321	11.5	79	3	US-09-027-376-14

28	321	11.5	79	3	US-09-094-192-14	Sequence 14, Appl
29	268.5	9.6	257	3	US-08-596-684F-7	Sequence 7, Appl
30	239.5	8.6	60	2	US-08-584-671-12	Sequence 12, Appl
31	239.5	8.6	60	3	US-09-027-376-12	Sequence 12, Appl
32	239.5	8.6	60	3	US-09-268-070-4	Sequence 4, Appl
33	239.5	8.6	60	3	US-09-094-192-12	Sequence 12, Appl
34	238.5	8.6	68	3	US-09-268-070-3	Sequence 3, Appl
35	208	7.5	67	3	US-09-268-070-3	Sequence 3, Appl
36	168.5	6.0	181	3	US-08-848-580-12	Sequence 12, Appl
37	168.5	6.0	181	4	US-08-488-123-12	Sequence 12, Appl
38	126.5	4.5	514	4	US-09-107-532A-6282	Sequence 6282, Ap
39	115.5	4.1	1002	4	US-09-762-724-4	Sequence 4, Appl
40	113	4.1	1027	4	US-09-762-724-4	Sequence 4, Appl
41	113	4.1	1786	3	US-08-973-462-8	Sequence 8, Appl
42	112.5	4.0	1014	4	US-09-762-724-2	Sequence 2, Appl
43	111	4.0	22	1	US-08-100-247-1	Sequence 1, Appl
44	111	4.0	22	1	US-08-483-146A-1	Sequence 1, Appl
45	111	4.0	22	1	US-08-483-146A-1	Sequence 11, Appl

ALIGNMENTS

```
RESULT 1
US-09-352-548-1
; Sequence 1, Application US/09352548
; Patent No. 6500431
; GENERAL INFORMATION:
; APPLICANT: Gill, Parkash S.
; APPLICANT: Parkash S. Gill, M.D., Inc.
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
; FILE REFERENCE: 017986-000410US
; CURRENT APPLICATION NUMBER: US/09/352, 548
; EARLIER FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: US 60/092, 647
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: prosaposin
; NAME/KEY: PEPTIDE
; LOCATION: (195)..(275)
; OTHER INFORMATION: Saposin B
US-09-352-548-1
Query Match 100.0%; Score 2789; DB 4; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.3e-242;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MYAFLPLASLIGALAGPVLGLKCTGSAVQCNVATASDCGAVKCLQTVNKKPTYS 60
Db 1 MYAFLPLASLIGALAGPVLGLKCTGSAVQCNVATASDCGAVKCLQTVNKKPTYS 60
QY 61 LPDICKDVVTAADMDLKDNTAEEELIVLEKTDMDLPPKPMASASCKEIVDSYLPVLDI 120
Db 61 LPDICKDVVTAADMDLKDNTAEEELIVLEKTDMDLPPKPMASASCKEIVDSYLPVLDI 120
QY 121 IKGEMSRGEVCSALNLCESIQKHLAEINHOKELENSKIPBLDTEVVAAPMAIPLLY 180
Db 121 IKGEMSRGEVCSALNLCESIQKHLAEINHOKELENSKIPBLDTEVVAAPMAIPLLY 180
QY 181 POGPRSKPOPKNDGVCODCIQVWTIDIQVTRNSTFVQALVBEKBCDRLPGMADI 240
Db 181 POGPRSKPOPKNDGVCODCIQVWTIDIQVTRNSTFVQALVBEKBCDRLPGMADI 240
QY 241 CKNTISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMQTLVPAKVASKNVIPAELVE 300
Db 241 CKNTISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMQTLVPAKVASKNVIPAELVE 300
```

QY 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMCKSKLPKSLSECCQEV 360
DB 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMCKSKLPKSLSECCQEV 360
QY 361 VDTYSSILSLILEEVSPELVCSMLHLCGSTRLPALTVHTOPKDGAFCEVCKLVGYLD 420
DB 361 VDTYSSILSLILEEVSPELVCSMLHLCGSTRLPALTVHTOPKDGAFCEVCKLVGYLD 420
QY 421 RNLEKSTKOEILAALEKGSFLPDYOKCQDFVAEYEPVLEILVEWMDPSFVCLKIG 480
DB 421 RNLEKSTKOEILAALEKGSFLPDYOKCQDFVAEYEPVLEILVEWMDPSFVCLKIG 480
QY 481 ACPSAHKPLGTGTEKCIWGPSYWCNTETAACNAVHCKRHWN 524
DB 481 ACPSAHKPLGTGTEKCIWGPSYWCNTETAACNAVHCKRHWN 524

RESULT 2
US-08-100-247-2
Sequence 2, Application US/08100247
Patent No. 5571787
GENERAL INFORMATION:
APPLICANT: O'BRIEN, JOHN S.
TITLE OF INVENTION: PROSAPOSIN AS A NEUTROTROPHIC FACTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/100,247
FILING DATE: 19930730
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: O'BRIEN.002A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: linear
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
IMMEDIATE SOURCE:
CLONE: PROSAPOSIN
US-08-100-247-2

Query Match 99.2%; Score 2766.5; DB 1; Length 523;
Best Local Similarity 99.4%; Pred. No. 1.4e-240;
Matches 523; Conservative 2; Mismatches 0; Indels 1; Gaps 1;
QY 1 MYAFELASLIGALAGPVGLKCTGSAVWCNVTTADCGAVHGLCTVNNKPTVVS 60
DB 1 MYAFELASLIGALAGPVGLKCTGSAVWCNVTTADCGAVHGLCTVNNKPTVVS 60
QY 61 LPDCICKDVTAAGMDLKDNTATEEELVYLEKTCMDLPRKPMASCKEIVDSYLPVILDI 120

DB 61 LPDCICKDVTAAGMDLKDNTATEEELVYLEKTCMDLPRKPMASCKEIVDSYLPVILDI 120
QY 121 IKGEMSRGECVSAALNICESEIQRKLAFLNOKOLESKITBELDTEVVAFPFMANIPALLY 180
DB 121 IKGEMSRGECVSAALNICESEIQRKLAFLNOKOLESKITBELDTEVVAFPFMANIPALLY 180
QY 181 POGDPRSKPOPKNDGVQCOCIDQWVTDIQTAVRTNSTFVQALVEHVEECDRIGPGMADI 240
DB 181 POGDPRSKPOPKNDGVQCOCIDQWVTDIQTAVRTNSTFVQALVEHVEECDRIGPGMADI 240
QY 241 CKNTYSQYSEIALQMMHMQPKELCALVGFCEVKEKMPQTLVPAKASNVIPALELVE 300
DB 240 CKNTYSQYSEIALQMMHMQPKELCALVGFCEVKEKMPQTLVPAKASNVIPALELVE 300
QY 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMCKSKLPKSLSECCQEV 360
DB 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMCKSKLPKSLSECCQEV 360
QY 361 VDTYSSILSLILEEVSPELVCSMLHLCGSTRLPALTVHTOPKDGAFCEVCKLVGYLD 420
DB 361 VDTYSSILSLILEEVSPELVCSMLHLCGSTRLPALTVHTOPKDGAFCEVCKLVGYLD 420
QY 421 RNLEKSTKOEILAALEKGSFLPDYOKCQDFVAEYEPVLEILVEWMDPSFVCLKIG 480
DB 421 RNLEKSTKOEILAALEKGSFLPDYOKCQDFVAEYEPVLEILVEWMDPSFVCLKIG 480
QY 481 ACPSAHKPLGTGTEKCIWGPSYWCNTETAACNAVHCKRHWN 524
DB 481 ACPSAHKPLGTGTEKCIWGPSYWCNTETAACNAVHCKRHWN 524

RESULT 3
US-08-756-031-2
Sequence 2, Application US/08756031
Patent No. 6590074
GENERAL INFORMATION:
APPLICANT: O'BRIEN, JOHN S.
TITLE OF INVENTION: PROSAPOSIN AS A NEUTROTROPHIC FACTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/756,031
FILING DATE: 26-NOV-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/100,247
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: O'BRIEN.002A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: linear

MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
IMMEDIATE SOURCE:
CLONE: PROSAPOSIN
US-08-756-031-2

Query Match 99.2%; Score 2766.5; DB 4; Length 523;
Best Local Similarity 99.4%; Pred. No. 1.4e-240;
Matches 521; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 MYALFLASLIGALAGVGLGKCTRGSAVWCQNVKTASDCGAVKHCLOTVMNKPVTYS 60
DB 1 MYALFLASLIGALAGVGLGKCTRGSAVWCQNVKTASDCGAVKHCLOTVMNKPVTYS 60
QY 61 LPDCICKDVTAAGDMLKDNATEEELVYLEKTCMDLPKPMNASCKEIVDSYLPVILI 120
DB 61 LPDCICKDVTAAGDMLKDNATEEELVYLEKTCMDLPKPMNASCKEIVDSYLPVILI 120
QY 121 IKGEMSRGGEVCSALNLCESLQKHLAEINHOQKLESNKIPELDMTEVVAPEPMANIPLLLY 180
DB 121 IKGEMSRGGEVCSALNLCESLQKHLAEINHOQKLESNKIPELDMTEVVAPEPMANIPLLLY 180
QY 181 PODGPRSKPOKNDGVCODCIQWVTDIQTAVRTNSTFVQALVEHKECDRLGPMADI 240
DB 181 PODGPRSKPOKNDGVCODCIQWVTDIQTAVRTNSTFVQALVEHKECDRLGPMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPKEICALVGFCDVEKEMQTLVPAKVASKNVIPALETVE 300
DB 240 CKNYISQYSEIAIQMMHMQPKEICALVGFCDVEKEMQTLVPAKVASKNVIPALETVE 300
QY 301 PIKKEVPAKSDVYCEVEFLVKEVTKLIDNNKTEKEILDADFDCSKLPSLSSECOEV 360
DB 301 PIKKEVPAKSDVYCEVEFLVKEVTKLIDNNKTEKEILDADFDCSKLPSLSSECOEV 360
QY 360 VDTYSSILSLIEEVSPELVCSMLHCSGTRLPALVTAVHTOPKDGGECEVCKLVGYLD 419
DB 360 VDTYSSILSLIEEVSPELVCSMLHCSGTRLPALVTAVHTOPKDGGECEVCKLVGYLD 419
QY 421 RNLEKNSIKQEIILALEKGCFLPDPPYKQCDQFAAEVEPVLIELVEMDPSFVCLKIG 480
DB 421 RNLEKNSIKQEIILALEKGCFLPDPPYKQCDQFAAEVEPVLIELVEMDPSFVCLKIG 480
QY 480 ACPSAHKPLGTGTEKICWGPSTWCONTEETAOCNAVEHCKRHVMN 523
DB 480 ACPSAHKPLGTGTEKICWGPSTWCONTEETAOCNAVEHCKRHVMN 523

RESULT 4
US-08-232-513A-3
Sequence 3, Application US/08232513A
Patent No. 5700909
GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides
TITLE OF INVENTION: as Therapeutic Agents
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSES: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,513A
FILING DATE: 21-APR-1994

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/100,247
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UD 1643
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9901
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..523
OTHER INFORMATION: /label= Hum_prosaposin
US-08-232-513A-3

Query Match 98.9%; Score 2759.5; DB 1; Length 523;
Best Local Similarity 99.2%; Pred. No. 5.9e-240;
Matches 520; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 MYALFLASLIGALAGVGLGKCTRGSAVWCQNVKTASDCGAVKHCLOTVMNKPVTYS 60
DB 1 MYALFLASLIGALAGVGLGKCTRGSAVWCQNVKTASDCGAVKHCLOTVMNKPVTYS 60
QY 61 LPDCICKDVTAAGDMLKDNATEEELVYLEKTCMDLPKPMNASCKEIVDSYLPVILI 120
DB 61 LPDCICKDVTAAGDMLKDNATEEELVYLEKTCMDLPKPMNASCKEIVDSYLPVILI 120
QY 121 IKGEMSRGGEVCSALNLCESLQKHLAEINHOQKLESNKIPELDMTEVVAPEPMANIPLLLY 180
DB 121 IKGEMSRGGEVCSALNLCESLQKHLAEINHOQKLESNKIPELDMTEVVAPEPMANIPLLLY 180
QY 181 PODGPRSKPOKNDGVCODCIQWVTDIQTAVRTNSTFVQALVEHKECDRLGPMADI 240
DB 181 PODGPRSKPOKNDGVCODCIQWVTDIQTAVRTNSTFVQALVEHKECDRLGPMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPKEICALVGFCDVEKEMQTLVPAKVASKNVIPALETVE 300
DB 241 CKNYISQYSEIAIQMMHMQPKEICALVGFCDVEKEMQTLVPAKVASKNVIPALETVE 300
QY 301 PIKKEVPAKSDVYCEVEFLVKEVTKLIDNNKTEKEILDADFDCSKLPSLSSECOEV 360
DB 301 PIKKEVPAKSDVYCEVEFLVKEVTKLIDNNKTEKEILDADFDCSKLPSLSSECOEV 360
QY 360 VDTYSSILSLIEEVSPELVCSMLHCSGTRLPALVTAVHTOPKDGGECEVCKLVGYLD 419
DB 360 VDTYSSILSLIEEVSPELVCSMLHCSGTRLPALVTAVHTOPKDGGECEVCKLVGYLD 419
QY 421 RNLEKNSIKQEIILALEKGCFLPDPPYKQCDQFAAEVEPVLIELVEMDPSFVCLKIG 480
DB 421 RNLEKNSIKQEIILALEKGCFLPDPPYKQCDQFAAEVEPVLIELVEMDPSFVCLKIG 480
QY 480 ACPSAHKPLGTGTEKICWGPSTWCONTEETAOCNAVEHCKRHVMN 524
DB 480 ACPSAHKPLGTGTEKICWGPSTWCONTEETAOCNAVEHCKRHVMN 523

RESULT 5
US-08-483-146A-2
Sequence 2, Application US/08483146A
Patent No. 5696080
GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED

```

; TITLE OF INVENTION: THEREFROM
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Blvd. 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,146A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelson, Ned A
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: MYELOS.002DV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
;
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
;
; US-08-483-146A-2
;
Query Match      98.9%; Score 2757.5; DB 1; Length 523;
Best Local Similarity 99.2%; Pred. No. 8.9e-240;
Matches 520; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
;
QY 1 MYALFLASLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVHCLQTVWVKPTYS 60
DB 1 MYALFLASLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVHCLQTVWVKPTYS 60
QY 61 LPDICKDVYTAAGDMVKONATEEETLVYLEKTCDWLPRPNMSASCKEIVDSYLPVILDI 120
DB 61 LPDICKDVYTAAGDMVKONATEEETLVYLEKTCDWLPRPNMSASCKEIVDSYLPVILDI 120
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQQLBSNKPILDMTEVVAAPFMANIPILLY 180
DB 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQQLBSNKPILDMTEVVAAPFMANIPILLY 180
QY 181 FQDGRSKPQPKD-GDVCODCIQWVTDIQTAVRTNSTFVQALVEHVEKECDRLGPGMADI 240
DB 181 FQDGRSKPQPKD-GDVCODCIQWVTDIQTAVRTNSTFVQALVEHVEKECDRLGPGMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPEIKALVGFCDVEKEMQTLVPKVASKNVIVPALEIVE 300
DB 241 CKNYISQYSEIAIQMMHMQPEIKALVGFCDVEKEMQTLVPKVASKNVIVPALEIVE 300
QY 301 PIKKEHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMKCSLPSLSBECQEV 360
DB 301 PIKKEHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMKCSLPSLSBECQEV 360
QY 361 VDTGSSSLTSLLEEVSELYCSMLHLCSGTRLPALTVHVTQPKDGGCEVCKKLVGLD 420
DB 361 VDTGSSSLTSLLEEVSELYCSMLHLCSGTRLPALTVHVTQPKDGGCEVCKKLVGLD 420
QY 421 RNLEKNSTKQETILALEKGSFLDPYQKOCDFVAIEVPEVLIELVWMDPSFVCLKIG 480
DB 421 RNLEKNSTKQETILALEKGSFLDPYQKOCDFVAIEVPEVLIELVWMDPSFVCLKIG 480

```

```

DB 420 RNLEKNSTKQETILALEKGSFLDPYQKOCDFVAIEVPEVLIELVWMDPSFVCLKIG 479
QY 481 ACPSAHKPRLGTEKICWGPSYWCNTETAAQCNVAHECKEHWNN 524
DB 480 ACPSAHKPRLGTEKICWGPSYWCNTETAAQCNVAHECKEHWNN 523
;
RESULT 6
US-08-484-594A-2
; Sequence 2, Application US/08484594A
; Patent No. 5714459
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; APPLICANT: Kishimoto, Yasuo
; TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUROTROPHIC PEPTIDES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,594A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/100,247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelson, Ned A
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: MYELOS.002DV2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
;
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
;
; US-08-484-594A-2
;
Query Match      98.9%; Score 2757.5; DB 1; Length 523;
Best Local Similarity 99.2%; Pred. No. 8.9e-240;
Matches 520; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
;
QY 1 MYALFLASLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVHCLQTVWVKPTYS 60
DB 1 MYALFLASLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVHCLQTVWVKPTYS 60
QY 61 LPDICKDVYTAAGDMVKONATEEETLVYLEKTCDWLPRPNMSASCKEIVDSYLPVILDI 120
DB 61 LPDICKDVYTAAGDMVKONATEEETLVYLEKTCDWLPRPNMSASCKEIVDSYLPVILDI 120
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQQLBSNKPILDMTEVVAAPFMANIPILLY 180
DB 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQQLBSNKPILDMTEVVAAPFMANIPILLY 180
QY 181 FQDGRSKPQPKD-GDVCODCIQWVTDIQTAVRTNSTFVQALVEHVEKECDRLGPGMADI 240
DB 181 FQDGRSKPQPKD-GDVCODCIQWVTDIQTAVRTNSTFVQALVEHVEKECDRLGPGMADI 240

```

QY 241 CKNYISQYSEIAIQMMHMOPEKICALVGFCDVEKEMPMQTLVPAKVASKNVIIPALDELVE 300
DB 240 CKNYISQYSEIAIQMMHMOPEKICALVGFCDVEKEMPMQTLVPAKVASKNVIIPALDELVE 299
QY 301 PIKKEHVPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAPDKKCSKLPKSLSECCOV 360
DB 300 PIKKEHVPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAPDKKCSKLPKSLSECCOV 359
QY 361 VDTYSSSLTSLILEEVSPELVCSMLHLCGSTRLPALTYHTVTPKDGGFCEVCKLVGYLD 420
DB 360 VDTYSSSLTSLILEEVSPELVCSMLHLCGSTRLPALTYHTVTPKDGGFCEVCKLVGYLD 419
QY 421 RNLEKNSTKOEILALAEKGCSEFLPDYKQCDQFVAEYEPVLIILVEMDPSFVCLKIG 480
DB 420 RNLEKNSTKOEILALAEKGCSEFLPDYKQCDQFVAEYEPVLIILVEMDPSFVCLKIG 479
QY 481 ACPSAHKPLGTETKCIWGPSTWCONTEPAQCNAAVEHCKRHVN 524
DB 480 ACPSAHKPLGTETKCIWGPSTWCONTEPAQCNAAVEHCKRHVN 523

RESULT 7
US-09-076-258A-2

; Sequence 2, Application US/09076258A
; Patent No. 6559124

GENERAL INFORMATION:

APPLICANT: O'Brien, John S.

APPLICANT: Kishimoto, Yasuo

TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS

TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUTROPHIC PEPTIDES DERIVED

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobb, Martens, Olson and Bear

STREET: 620 Newport Center Blvd. 16th Floor

CITY: Newport Beach

STATE: CA

COUNTRY: USA

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/076,258A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/958,970

FILING DATE: 28-OCT-97

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/483,146

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: 08/100,247

FILING DATE: 30-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: Israel, Ned A

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: MEYLOS,2DV1C2

TELEPHONE: 619-235-8550

TELEFAX: 619-235-0176

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 523 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: N-terminal

US-09-076-258A-2

Query Match 98.9%; Score 2757.5; DB 4; Length 523;

Best Local Similarity 99.2%; Pred. No. 8.9e-240;

Matches 520; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 MYALFLASLIGALAGVPLGLKECTRGSAVWQNVKPAASDCAVHCLQTVNKPVTYS 60
DB 1 MYALFLASLIGALAGVPLGLKECTRGSAVWQNVKPAASDCAVHCLQTVNKPVTYS 60
QY 61 LPDCIKDQVTAAGDMLKDNATEEILVYLEKTCOMLPKPMNASCKEIVDSYLPVILI 120
DB 61 LPDCIKDQVTAAGDMLKDNATEEILVYLEKTCOMLPKPMNASCKEIVDSYLPVILI 120
QY 121 IKGEMSRGCVCSALNLCESIQKHLAEINHQKLESNKIPBLDMTEVVAFFPMANIPLLY 180
DB 121 IKGEMSRGCVCSALNLCESIQKHLAEINHQKLESNKIPBLDMTEVVAFFPMANIPLLY 180
QY 181 PDGPRSRKPPKNDNDVDCDQIOMVTDIQTAVRTNSTFVQALVEHVEKCDRLGPMADI 240
DB 181 PDGPRSRKPPKNDNDVDCDQIOMVTDIQTAVRTNSTFVQALVEHVEKCDRLGPMADI 239
QY 241 CKNYISQYSEIAIQMMHMOPEKICALVGFCDVEKEMPMQTLVPAKVASKNVIIPALDELVE 300
DB 240 CKNYISQYSEIAIQMMHMOPEKICALVGFCDVEKEMPMQTLVPAKVASKNVIIPALDELVE 299
QY 301 PIKKEHVPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAPDKKCSKLPKSLSECCOV 360
DB 301 PIKKEHVPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAPDKKCSKLPKSLSECCOV 359
QY 361 VDTYSSSLTSLILEEVSPELVCSMLHLCGSTRLPALTYHTVTPKDGGFCEVCKLVGYLD 420
DB 360 VDTYSSSLTSLILEEVSPELVCSMLHLCGSTRLPALTYHTVTPKDGGFCEVCKLVGYLD 419
QY 421 RNLEKNSTKOEILALAEKGCSEFLPDYKQCDQFVAEYEPVLIILVEMDPSFVCLKIG 480
DB 420 RNLEKNSTKOEILALAEKGCSEFLPDYKQCDQFVAEYEPVLIILVEMDPSFVCLKIG 479
QY 481 ACPSAHKPLGTETKCIWGPSTWCONTEPAQCNAAVEHCKRHVN 524
DB 480 ACPSAHKPLGTETKCIWGPSTWCONTEPAQCNAAVEHCKRHVN 523

RESULT 8

US-09-352-548-2

; Sequence 2, Application US/09352548

; Patent No. 6500431

GENERAL INFORMATION:

APPLICANT: Parkash S. Gill, M.D., Inc.

TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth

FILE REFERENCE: 01/986-000410US

CURRENT APPLICATION NUMBER: US/09/352,548

CURRENT FILING DATE: 1999-07-12

EARLIER APPLICATION NUMBER: US 60/092,647

EARLIER FILING DATE: 1998-07-13

NUMBER OF SEQ ID NOS: 59

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 81

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Saposin B

US-09-352-548-2

Query Match 15.5%; Score 431; DB 4; Length 81;

Best Local Similarity 100.0%; Pred. No. 1.7e-31;

Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 GDVCCDQIOMVTDIQTAVRTNSTFVQALVEHVEKCDRLGPMADICKNYISQYSEIAIQ 254
DB 1 GDVCCDQIOMVTDIQTAVRTNSTFVQALVEHVEKCDRLGPMADICKNYISQYSEIAIQ 60

OY 255 MMHMPKICALVGFCDK 275
Db 61 MMHMPKICALVGFCDK 81

RESULT 9

US-08-584-671-15
Sequence 15, Application US/08584671
Patent No. 5910568
GENERAL INFORMATION:
APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM
TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA
STATE UNIVERSITY
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/584,671
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MONAHAN, THOMAS J
REGISTRATION NUMBER: 29835
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 814-865-6277
TELEFAX: 814-865-3591
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 80
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: UNKNOWN
US-08-584-671-15

Query Match 15.0%; Score 418.5; DB 2; Length 80;
Best Local Similarity 98.8%; Pred. No. 2.2e-30;

Matches 80; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 128 PGEVCSALNLCESLQKHLAEINHOQLESNKIPELDMEVAPPMANIPLLIYPODGRS 187
Db 1 PGEVC-ALNLCESLQKHLAEINHOQLESNKIPELDMEVAPPMANIPLLIYPODGRS 59
OY 188 KPQKNDGVCDCCIOGMTDI 208
Db 60 KPQKNDGVCDCCIOGMTDI 80

RESULT 10

US-09-027-376-15
Sequence 15, Application US/09027376
Patent No. 6004586
GENERAL INFORMATION:
APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM
TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE
NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA
STATE UNIVERSITY
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/027,376
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/584,671
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MONAHAN, THOMAS J
REGISTRATION NUMBER: 29835
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 814-865-6277
TELEFAX: 814-865-3591
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 80
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: UNKNOWN
US-09-027-376-15

Query Match 15.0%; Score 418.5; DB 3; Length 80;
Best Local Similarity 98.8%; Pred. No. 2.2e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 128 PGEVCSALNLCESLQKHLAEINHOQLESNKIPELDMEVAPPMANIPLLIYPODGRS 187
Db 1 PGEVC-ALNLCESLQKHLAEINHOQLESNKIPELDMEVAPPMANIPLLIYPODGRS 59
OY 188 KPQKNDGVCDCCIOGMTDI 208
Db 60 KPQKNDGVCDCCIOGMTDI 80

RESULT 11

US-09-094-192-15
Sequence 15, Application US/09094192
Patent No. 6103483
GENERAL INFORMATION:
APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM TO EGG SURFACES AND PROCI
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA STATE UNIVERSITY
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/094,192
FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: MONAHAN, THOMAS J
REGISTRATION NUMBER: 29835
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 814-865-6277
TELEFAX: 814-865-3591
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 80
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: UNKNOWN
US-09-094-192-15

Query Match 14.9%; Score 415.5; DB 3; Length 80;
Best Local Similarity 97.5%; Pred. No. 4.1e-30;
Matches 79; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 128 PGEVSAINTLCEISLOKHIAELNHOKLESNKIPELDMTEVVAPEPMANIPILLYPODGPBS 187
Db 1 PGEVCA-ALNLCOSLOKHIAELNHOKLESNKIPELDMTEVVAPEPMANIPILLYPODGPBS 59
Qy 188 KPOPKNGDVCCDCIOWTDTI 208
Db 60 KPOPKNGDVCCDCIOWTDTI 80

RESULT 12
US-08-100-247-3
Sequence 3, Application US/08100247
Patent No. 5571787

GENERAL INFORMATION:
APPLICANT: O'BRIEN, JOHN S.
APPLICANT: KISHIMOTO, YASUO
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/100.247
FILING DATE: 19930730

CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: OBRIEN.002A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
IMMEDIATE SOURCE:
CLONE: SAPOSIN C
US-08-100-247-3

Query Match 14.8%; Score 412; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 SDVYCEVEFLVKEVTKLIDNNKTEKELIDAFDMCKSLPKSLSECEQVVDYGSSTLS 370
Db 1 SDVYCEVEFLVKEVTKLIDNNKTEKELIDAFDMCKSLPKSLSECEQVVDYGSSTLS 60

Qy 371 ILLEEVSPELVCSMLHLCSG 390
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 13
US-08-483-146A-3
Sequence 3, Application US/08483146A
Patent No. 5696080

GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
APPLICANT: Kishimoto, Yasuo
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED
FROM THEREFROM
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Blvd. 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483.146A
FILING DATE: 07-JUN-1995

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.002DV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-483-146A-3

Query Match 14.8%; Score 412; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.5e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 SDVYCEVEFLVKEVTKLIDNNKTEKELIDAFDMCKSLPKSLSECEQVVDYGSSTLS 370
Db 1 SDVYCEVEFLVKEVTKLIDNNKTEKELIDAFDMCKSLPKSLSECEQVVDYGSSTLS 60

Qy 371 ILLEEVSPELVCSMLHLCSG 390
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 14

US-08-232-513A-4
Sequence 4, Application US/08232513A
Patent No. 5700909
GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides
TITLE OF INVENTION: as Therapeutic Agents
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,513A
FILING DATE: 21-Apr-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/100,247
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UD 1643
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..80
OTHER INFORMATION: /label= Saposin_C
US-08-232-513A-4

Query Match

Best Local Similarity 14.8%; Score 412; DB 1; Length 80;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMKSKLPKSLSECCQEVVDYSSILS 370
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMKSKLPKSLSECCQEVVDYSSILS 60
OY 371 ILLEVSPELVCSMLHLCSG 390
Db 61 ILLEVSPELVCSMLHLCSG 80

RESULT 15

US-08-484-594A-3
Sequence 3, Application US/08484594A
Patent No. 5714459
GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
APPLICANT: Kishimoto, Yasuo
TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUTROTROPHIC PEPTIDES
TITLE OF INVENTION: DERIVED THEREFROM
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,594A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/100,247
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.002DV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-484-594A-3

Query Match

Best Local Similarity 14.8%; Score 412; DB 1; Length 80;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMKSKLPKSLSECCQEVVDYSSILS 370
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMKSKLPKSLSECCQEVVDYSSILS 60
OY 371 ILLEVSPELVCSMLHLCSG 390
Db 61 ILLEVSPELVCSMLHLCSG 80

Search completed: May 5, 2004, 13:16:43
Job time : 23.8908 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:15:48 ; Search time 73.151 Seconds

(without alignments)
2023.963 Million cell updates/sec

Title: US-09-743-684A-1

Sequence: 1 MYALFLIASLGAALAGPVL.....NTETRAQCNVHECKRWVN 524

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2789	100.0	524	2	AAR70783 Prosaposin
2	2789	100.0	524	2	AAW85652 Human pro
3	2789	100.0	524	3	AAV58716 Human pro
4	2789	100.0	524	6	ABU79099 Lip-TAA b
5	2789	100.0	524	6	ABU05200 Human exp
6	2789	100.0	524	6	ABU05207 Human exp
7	2789	100.0	524	6	ABU05203 Human exp
8	2789	100.0	524	6	ABU07340 Human exp
9	2789	100.0	524	6	ABU05216 Human exp
10	2789	100.0	524	6	ABU05202 Human exp
11	2789	100.0	524	6	ABU05214 Human exp
12	2789	100.0	524	6	ABU05215 Human exp
13	2789	100.0	524	6	ABU05199 Human exp
14	2789	100.0	524	6	ABU05212 Human exp
15	2789	100.0	524	6	ABU05213 Human exp
16	2789	100.0	524	6	ABU05205 Human exp
17	2779	99.6	527	5	ABU05208 Human exp
18	2777.5	99.6	527	5	AAAB1915 Amino aci
19	2777.5	99.6	527	5	ABP68602 Human pan
20	2777.5	99.6	527	6	ABU79100 Lip-TAA b
21	2777.5	99.6	527	6	ABU05204 Human exp
22	2777.5	99.6	527	6	ABU05210 Human exp
23	2772.5	99.4	523	4	AAAB1916 Amino aci
24	2772.5	99.4	523	6	ABU05211 Human exp
25	2768	99.2	526	6	ABU05209 Human exp

26	2767.5	99.2	527	6	ABU05206 Human exp
27	2449.5	87.8	479	6	ABR39442 Human GEN
28	2024.5	72.6	385	6	ABR41750 Human DIT
29	1969	70.6	354	7	ADB85295 Rat tubul
30	1921	68.9	554	5	ABB57102 Mouse 1sc
31	1201	43.1	521	3	AAV56966 Human SBP
32	1201	43.1	521	4	AAW78587 Human pro
33	1201	43.1	521	4	AAW79571 Human pro
34	1201	43.1	521	4	AAAB86362 Human pro
35	1201	43.1	521	4	AAE15547 Human sec
36	1196	42.9	521	4	AAW79572 Human pro
37	1139.5	40.9	531	5	AAE01770 Human gen
38	1139.5	40.9	531	5	ABG64156 Human alb
39	1125	40.3	210	6	ABU05201 Human exp
40	1093	39.2	209	5	ABG70166 Human pre
41	834.5	29.9	362	4	AAW78588 Human pro
42	834.5	29.9	362	5	ABB07499 Human lip
43	812	29.1	153	6	ABU70504 Human adi
44	811	29.1	153	6	ABU70799 Human adi
45	676	24.2	129	6	ABU70422 Human adi

ALIGNMENTS

RESULT 1
AAR70783
ID AAR70783 standard; protein; 524 AA.
XX
XX AAR70783;
AC
XX
DT 25-MAR-2003 (revised)
DT 30-AUG-1995 (first entry)
XX
XX DE Prosaposin.
XX
XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;
KW neurotrophic peptide; multiple sclerosis; leukoencephalitis;
KW adrenal leukodystrophy; prosaposin.
XX
XX Homo sapiens.
OS
XX
XX W09503821-A1.
XX
XX 09-FEB-1995.
PD
XX
XX 28-JUL-1994; 94WO-US008453.
PF
XX
XX 30-JUL-1993; 93US-00100247.
PR 21-APR-1994; 94US-00232513.
XX
XX (OBRI/) O'BRIEN J S.
PA
XX
PI O'Brien JS, Kishimoto Y;
XX
XX WPI; 1995-082029/11.
DR N-PSDB; AAQ85355.
XX
XX Stimulating neural cell out-growth and myelination - with pro:saposin,
PT saposin C or new neurotrophic peptide(s) from cytokine(s), for treating
PT nervous system diseases.
XX
XX
XX Disclosure; Page 30-32; 50pp; English.
PS
XX
XX The peptide given in AAR70773, corresponding to amino acids 8-29 of human
CC saposin-C (AAR70784), promotes neurite outgrowth in vitro. A consensus
CC sequence was determined by comparing the peptide with hematopoietic and
CC neurotrophic cytokines, and neurotrophic peptides (AAR70774-82) were
CC identified in the AB loop of human ciliary neurotrophic factor,
CC interleukins-6, -2, -3 and -gamma, erythropoietin and leukocyte
CC inhibitory factor, and in helix C of human interleukin-1-beta and
CC oncostatin-M. Prosaposin (AAR70783) and saposin-C also promoted nerve
CC cell myelination ex vivo. (Updated on 25-MAR-2003 to correct PN field.)

CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to
 CC correct PI field.)

CC Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 2; Length 524;
 Best Local Similarity 100.0%; Pred. No. 2,3e-237;
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MYALFLASLIGALAGPVLGKECTRSAYWCQNVKTASDCGAVHCLQTVWNPPTVKS 60
DB 1 MYALFLASLIGALAGPVLGKECTRSAYWCQNVKTASDCGAVHCLQTVWNPPTVKS 60
QY 61 LPEDICKDVVTAAGDMLKDNATEEELVYLEKTCMDLPKPNMSASCKEIVDSYLVILDI 120
DB 61 LPEDICKDVVTAAGDMLKDNATEEELVYLEKTCMDLPKPNMSASCKEIVDSYLVILDI 120
QY 121 IGENSRPGEVCSALNLCESIQKHAEINHQKQLESNNKIPELDMTEVVAPEMANIPLLLY 180
DB 121 IGENSRPGEVCSALNLCESIQKHAEINHQKQLESNNKIPELDMTEVVAPEMANIPLLLY 180
QY 181 PDGGRSRKQPKNDGVCODCIQWVTDIQTAVRNSFTVQALVHVKEECRLPGMADI 240
DB 181 PDGGRSRKQPKNDGVCODCIQWVTDIQTAVRNSFTVQALVHVKEECRLPGMADI 240
QY 241 CKNYISQYSEIALQMMHMQPEKICALVGFCEVKEPMQTLVPAKVASKNVIPALIEVE 300
DB 241 CKNYISQYSEIALQMMHMQPEKICALVGFCEVKEPMQTLVPAKVASKNVIPALIEVE 300
QY 301 PIKKEHVPAKSDVYCEVEFLVKEVTYKLIIDNNKTEKEILDAPDKMCSKLPKSLSECEQEV 360
DB 301 PIKKEHVPAKSDVYCEVEFLVKEVTYKLIIDNNKTEKEILDAPDKMCSKLPKSLSECEQEV 360
QY 361 VDTYSSILSLILEVSEPELVCSMLHSGTRLPALTVHVOQPKDGGCEVCCKLVGLD 420
DB 361 VDTYSSILSLILEVSEPELVCSMLHSGTRLPALTVHVOQPKDGGCEVCCKLVGLD 420
QY 421 RNLEKNSTKQEBILALEKSGSFLPDPYQKCDQFAVEYEPVILIEVMPSPFVCLXIG 480
DB 421 RNLEKNSTKQEBILALEKSGSFLPDPYQKCDQFAVEYEPVILIEVMPSPFVCLXIG 480
QY 481 ACPSAHKLPTGTEKICWGPSTWCONTEFAACQNAVEHCKRHWNV 524
DB 481 ACPSAHKLPTGTEKICWGPSTWCONTEFAACQNAVEHCKRHWNV 524

```

RESULT 2
 AAW85652
 ID AAW85652 standard; protein; 524 AA.

XX AAW85652;

XX 19-JUL-1999 (first entry)

XX Human prosaposin N-terminal peptide.

XX Prosaposin; saposin; prosapetides; prosaposin receptor agonists; PRA;
 XX peripheral nervous system; central nervous system; PNS; CNS; Akt; Bcl-2;
 XX therapy; treatment; apoptosis; caspase; tumor necrosis factor; TNF;
 XX cytokine; interferon gamma; IFN; inflammation; rheumatoid arthritis;
 XX Crohn's disease; irritable bowel syndrome; ashma; cardiac infarction;
 XX congestive heart failure; multiple sclerosis;
 XX acute disseminated inflammatory leukoencephalitis;
 XX progressive multifocal leukoencephalitis; Alzheimer's disease;
 XX Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;
 XX ischemic heart disease; Guillain-Barre disease; alopecia; AIDS dementia;
 XX cerebral malaria; HTLV; neuropathy;
 XX inflammatory neurodegenerative disease; toxin-induced liver disease.

XX Homo sapiens.

XX MO9912559-A1.

PD 18-MAR-1999.

XX 09-SEP-1998; 98WO-US019216.

XX 09-SEP-1997; 97US-0058352P.

XX 04-JUN-1998; 98US-0088129P.

XX (REGC) UNIV CALIFORNIA.

XX O'Brien US;

XX WPI; 1999-228139/19.

XX N-PSDB; AAX08488.

XX Use of prosaposin receptor agonist.

XX Claim 7; Fig 2; 90pp; English.

Prosaposin is a 70kDa glycoprotein which is proteolytically processed to generate saposins A, B, C and D, all of which are similar to each other and have a similar placement of six cysteines, a glycosylation site and conserved proline residues. Prosaposin, saposin C and prosaposin derived peptides (prosapetides), have therapeutic applications in promoting recovery after toxic, traumatic, myocardial ischemic, degenerative and inherited lesions to the peripheral and central nervous system. Prosaposin receptor agonists (PRAs) inhibit proinflammatory cytokine-induced apoptosis by activation of the Ser/Thr protein kinase Akt. Akt dissociates complexes of Bcl-2 family members, such as BAD-Bcl-2, releasing Bcl-2 and its family members which inhibit caspases, thereby inhibiting apoptosis. An additional mechanism whereby PRAs inhibit apoptosis is by blocking activation of JNK, a proapoptotic signaling component. Within several minutes after binding to the receptor, PRAs block JNK activation induced by tumor necrosis factor- α (TNF α). The activation of JNK by TNF α is another well known mechanism for TNF α -induced, as well as other proinflammatory cytokine-induced apoptosis. The method can be used for inhibiting apoptosis which is caspase-mediated or induced by a proinflammatory cytokine which is TNF α or interferon- γ . It can be used for inhibiting apoptosis associated with a disorder such as e.g. rheumatoid arthritis, Crohn's disease, irritable bowel syndrome, ashma, cardiac infarction, congestive heart failure, multiple sclerosis, acute disseminated inflammatory leukoencephalitis, progressive multifocal leukoencephalitis, Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's disease, ischemic heart disease, Guillain-Barre disease, traumatic brain injury, traumatic spinal cord injury, alopecia, AIDS dementia, cerebral malaria, HTLV, neuropathy, inflammatory neurodegenerative disease, and toxin-induced liver disease. This 524 N-terminal peptide of prosaposin also acts as a prosaposin receptor agonist

XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 2; Length 524;
 Best Local Similarity 100.0%; Pred. No. 2,3e-237;
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MYALFLASLIGALAGPVLGKECTRSAYWCQNVKTASDCGAVHCLQTVWNPPTVKS 60
DB 1 MYALFLASLIGALAGPVLGKECTRSAYWCQNVKTASDCGAVHCLQTVWNPPTVKS 60
QY 61 LPEDICKDVVTAAGDMLKDNATEEELVYLEKTCMDLPKPNMSASCKEIVDSYLVILDI 120
DB 61 LPEDICKDVVTAAGDMLKDNATEEELVYLEKTCMDLPKPNMSASCKEIVDSYLVILDI 120
QY 121 IGENSRPGEVCSALNLCESIQKHAEINHQKQLESNNKIPELDMTEVVAPEMANIPLLLY 180
DB 121 IGENSRPGEVCSALNLCESIQKHAEINHQKQLESNNKIPELDMTEVVAPEMANIPLLLY 180
QY 181 PDGGRSRKQPKNDGVCODCIQWVTDIQTAVRNSFTVQALVHVKEECRLPGMADI 240
DB 181 PDGGRSRKQPKNDGVCODCIQWVTDIQTAVRNSFTVQALVHVKEECRLPGMADI 240
QY 241 CKNYISQYSEIALQMMHMQPEKICALVGFCEVKEPMQTLVPAKVASKNVIPALIEVE 300
DB 241 CKNYISQYSEIALQMMHMQPEKICALVGFCEVKEPMQTLVPAKVASKNVIPALIEVE 300

```


Db 241 CKNYISQYSEIAIQMMHMQPEKICALVGFCDVEKEMPMQTLVPAKASKNIVPALBELVE 300
 QY 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCLPRLSSECCOE 360
 Db 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCLPRLSSECCOE 360
 QY 361 VDTYGSSTLILBEVSPBELVCSMLHLCSGTRLPALTVHTVQPKDGGCEVCKLVGYLD 420
 Db 361 VDTYGSSTLILBEVSPBELVCSMLHLCSGTRLPALTVHTVQPKDGGCEVCKLVGYLD 420
 QY 421 RNLEKNSKQEIILALEKSGSFLPDPYKQCDQFVAEYEPVLIELVEMDPSFVCLKIG 480
 Db 421 RNLEKNSKQEIILALEKSGSFLPDPYKQCDQFVAEYEPVLIELVEMDPSFVCLKIG 480
 QY 481 ACPSAHKPLGTETKCIWGPSYWCNTETAAOCNAVEHCKRHVMN 524
 Db 481 ACPSAHKPLGTETKCIWGPSYWCNTETAAOCNAVEHCKRHVMN 524
 RESULT 3
 ID AAY58716 standard; protein, 524 AA.
 AC AAY58716;
 XX 25-APR-2000 (first entry)
 DT 25-APR-2000 (first entry)
 XX Human prosaposin.
 DB Human prosaposin.
 XX Prosaposin; saposin B; antiangiogenic; angiogenesis inhibitor;
 KM antitumour; antiproliferative; antimitratory; Kaposi's sarcoma; tumour;
 XX human; therapy.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH 195..275
 FT Protein /note="mature saposin B"
 FT 195..205
 FT Peptide /note="specifically claimed antiangiogenic peptide of
 Claim 23"
 FT 196..200
 FT Peptide /note="specifically claimed antiangiogenic peptide of
 Claim 4"
 FT 196..200
 XX WO200002902-A1.
 XX 20-JAN-2000.
 PD 20-JAN-2000.
 XX 12-JUL-1999; 99WO-US015772.
 PF 12-JUL-1999; 99WO-US015772.
 XX 13-JUL-1998; 98US-0092647P.
 PR 13-JUL-1998; 98US-0092647P.
 XX (GILL/) GILL P S.
 PA GILL PS;
 PI GILL PS;
 XX WPI: 2000-171128/15.
 DR WPI: 2000-171128/15.
 PT Saposin B derived peptides, useful as inhibitors of angiogenesis and
 PT tumor growth.
 XX Saposin B derived peptides, useful as inhibitors of angiogenesis and
 PS tumor growth.
 XX Disclosure; Page 18; 78pp; English.
 CC The present sequence is that of human prosaposin, a precursor of saposin
 CC B. The invention is based on the discovery that saposin B, previously
 CC known to be involved in the hydrolysis of sphingolipids, has potent
 CC antiangiogenic and antitumour activity, and also has antiproliferative
 CC and antimigratory activity against endothelial cells. This activity is
 CC conserved in cryptic polypeptides as small as 5 amino acids (see AAY58684
 CC -715), which can be synthetically prepared and used in vitro or in vivo
 CC for the treatment of undesired angiogenesis and tumor growth, especially
 CC Kaposi's sarcoma (claimed). The polypeptides can also be used in

CC conjunction with cytotoxic moieties to selectively kill certain cell
 CC types, e.g. for treatment of cancer, angiofibroma, neovascular glaucoma,
 CC arteriovenous malformation, nonunion fracture, arthritis and other
 CC connective tissue disorders, Osler-Weber syndrome, atherosclerotic
 CC plaque, psoriasis, corneal graft neovascularization, pyogenic granuloma,
 CC retrolental fibroplasia, diabetic retinopathy, scleroderma, haemangioma,
 CC trachoma, vascular adhesions and hypertrophic scars
 SQ Sequence 524 AA;
 Query Match 100.0%; Score 2789; DB 3; Length 524;
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MYALFLASLIGALAGPVGLKCTGSAWCONVTASCGAVKCLQTVNKKPYKS 60
 Db 1 MYALFLASLIGALAGPVGLKCTGSAWCONVTASCGAVKCLQTVNKKPYKS 60
 QY 61 LPDICKDVTAAGDMLKDNATEEELVLEKTCQDMLPKPMNSASCKEIVDSYPLVLDI 120
 Db 61 LPDICKDVTAAGDMLKDNATEEELVLEKTCQDMLPKPMNSASCKEIVDSYPLVLDI 120
 QY 121 IKGMSRGEVCSALNLCESLQKHLABELNHQKQLESNKIPELDTEVVAPEMANIPILLY 180
 Db 121 IKGMSRGEVCSALNLCESLQKHLABELNHQKQLESNKIPELDTEVVAPEMANIPILLY 180
 QY 181 PDDGPRSKPQPKDNGDYCCDCTQWVTDIQTAVRNTSTFVQALVBEKCECRLEGMAADI 240
 Db 181 PDDGPRSKPQPKDNGDYCCDCTQWVTDIQTAVRNTSTFVQALVBEKCECRLEGMAADI 240
 QY 241 CKNYISQYSEIAIQMMHMQPEKICALVGFCDVEKEMPMQTLVPAKASKNIVPALBELVE 300
 Db 241 CKNYISQYSEIAIQMMHMQPEKICALVGFCDVEKEMPMQTLVPAKASKNIVPALBELVE 300
 QY 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCLPRLSSECCOE 360
 Db 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCLPRLSSECCOE 360
 QY 361 VDTYGSSTLILBEVSPBELVCSMLHLCSGTRLPALTVHTVQPKDGGCEVCKLVGYLD 420
 Db 361 VDTYGSSTLILBEVSPBELVCSMLHLCSGTRLPALTVHTVQPKDGGCEVCKLVGYLD 420
 QY 421 RNLEKNSKQEIILALEKSGSFLPDPYKQCDQFVAEYEPVLIELVEMDPSFVCLKIG 480
 Db 421 RNLEKNSKQEIILALEKSGSFLPDPYKQCDQFVAEYEPVLIELVEMDPSFVCLKIG 480
 QY 481 ACPSAHKPLGTETKCIWGPSYWCNTETAAOCNAVEHCKRHVMN 524
 Db 481 ACPSAHKPLGTETKCIWGPSYWCNTETAAOCNAVEHCKRHVMN 524
 RESULT 4
 ID ABU79099 standard; protein, 524 AA.
 ABU79099
 XX ABU79099;
 AC ABU79099;
 XX 18-JUN-2003 (first entry)
 DT 18-JUN-2003 (first entry)
 XX Lip-TAA binding protein, Prosaposin.
 DE Lip-TAA binding protein, Prosaposin.
 XX Superantigen; SAg; staphylococcal enterotoxin; tumour; cancer; apoptosis;
 KM gene therapy; mammalian cell receptor; cytostatic;
 KM tumour associated lipid; energy; T cell; antigen presenting cell; APC;
 XX tumouricidal immunocyte; antitumour.
 OS Undentified.
 XX US2002177551-A1.
 XX 28-NOV-2002.
 PD 28-NOV-2002.
 XX 30-MAY-2001; 2001US-00870759.

XX 31-MAY-2000; 2000US-0208128P.
 PR (TERM/) TERNAN D S.
 XX Terman DS;
 PI WPI; 2003-361759/34.
 DR A mammalian cell receptor, useful in the treatment of cancer by binding
 XX to tumor associated lipids where the binding induces anergy or apoptosis
 PT in T cells and antigen presenting cells.
 XX Disclosure; Page; 167pp; English.
 PS
 XX The invention relates to a mammalian cell receptor, useful in the
 CC treatment of cancer, which binds to tumor associated lipids and induces
 CC anergy or apoptosis in the T cells and antigen presenting cells (APCs).
 CC Also included are a mammalian cell useful in the treatment of cancer
 CC where the receptor which binds tumor associated lipids and induces
 CC cellular inactivation or death is deleted or functionally deactivated,
 CC producing (M1) a tumouricidal immunocyte population in vivo in a mammal
 CC (by allowing tumour associated lipids to contact immunocytes in which
 CC receptors for immunosuppressive fatty acids, ceramides, glycolipids,
 CC sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides,
 CC sialylated glycans, lipopeptides and proteoglycolipids are inactivated or
 CC deleted), a construct useful in the treatment of cancer comprising a
 CC superantigen (SAG) nucleotide inserted into a virus, a mammalian T cell
 CC useful in the treatment of cancer (where an adaptor protein which
 CC inhibits T cell activation by tumour associated antigens is deleted or
 CC functionally deactivated), a composition useful in the treatment of
 CC cancer (comprising a lipid raft conjugated to a superantigen), producing
 CC (M2) a tumouricidal immunocyte population ex vivo in a mammal (by
 CC allowing tumour associated lipids to contact immunocytes, in which
 CC receptors for the lipids are inactivated or deleted to produce a
 CC tumouricidal immunocyte population, and administering the tumouricidal
 CC activated immunocyte to the host), producing (M3) a tumouricidal APC
 CC population ex vivo in a mammal (by allowing a tumour associated lipid to
 CC contact APCs, in which receptors for the tumour associated lipids are
 CC inactivated or deleted to produce a tumouricidally activated population,
 CC and administering APCs to the host), producing a tumouricidal T cell
 CC population ex vivo in a mammal (by allowing a tumour associated lipid to
 CC contact T cells, in which adaptor proteins, which inhibit T cell
 CC activation by tumour associated antigens, are deleted or functionally
 CC deactivated to produce a tumouricidal population of T cells, and
 CC administering the tumouricidally activated T cells to the host, or
 CC allowing a superantigen-lipid raft to contact T cells ex vivo, and
 CC administering the tumouricidally activated T cells to the host), treating
 CC (M5) cancer in a mammal (by administering a lipid binding molecule which
 CC binds immunosuppressive tumour associated lipids in vivo), producing (M6)
 CC a tumouricidal T cell population in vivo in a mammal (by allowing a
 CC tumour associated antigen to contact immunocytes in which adaptor
 CC proteins which inhibit T cell activation by tumour associated antigens
 CC are deleted or functionally deactivated) and producing (M7) a
 CC tumouricidal T cell population ex vivo in a mammal comprising allowing a
 CC superantigen-lipid raft conjugate to contact immunocytes in vivo. The
 CC receptors, methods and compositions are useful for treating cancers and
 CC tumours. Bacterial superantigens are co-administered or administered as
 CC fusion constructs with anti-tumour proteins or motifs. The present
 CC sequence represents a tumour antigen or a motif identifying a tumour
 CC antigen, which can be functionally deactivated in the method of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format from the
 CC US patent office website at
 CC "seqdata.uspto.gov/sequence.html?DocID=20020177551"
 CC
 XX
 XX Sequence 524 AA:

Query Match 100.0%; Score 2789; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 2,3e-237;
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MYALFLASLILGALAGVVLGLKECTRGSAVWCQNVKTASDCGAVHCHCLOTWNNKPTVKS 60

Db
 1 MYALFLASLILGALAGVVLGLKECTRGSAVWCQNVKTASDCGAVHCHCLOTWNNKPTVKS 60
 QY LPDLCIDVVTAAAGDMLKDNATEBEELVLEKTCMDLPRKNMSASCKEIVDSYLPVILDT 120
 Db LPDLCIDVVTAAAGDMLKDNATEBEELVLEKTCMDLPRKNMSASCKEIVDSYLPVILDT 120
 QY IKGEMSRGVEYCSALNTCESLQKHAEINHOKESENKIPELMTEVAVAFPMNIPILLY 180
 Db IKGEMSRGVEYCSALNTCESLQKHAEINHOKESENKIPELMTEVAVAFPMNIPILLY 180
 QY 121 IKGEMSRGVEYCSALNTCESLQKHAEINHOKESENKIPELMTEVAVAFPMNIPILLY 180
 Db 121 IKGEMSRGVEYCSALNTCESLQKHAEINHOKESENKIPELMTEVAVAFPMNIPILLY 180
 QY 181 PODEPRSKPQPKXGNDVCDQCICQWVTDIQTAVRTNSTFVQALVEHYEBCDRIGPGWADI 240
 Db 181 PODEPRSKPQPKXGNDVCDQCICQWVTDIQTAVRTNSTFVQALVEHYEBCDRIGPGWADI 240
 QY 241 CKNYISQYSEIATQMMHMQPKKICLVGFCDEKEMPTQTLVPAVASKNVPALELVE 300
 Db 241 CKNYISQYSEIATQMMHMQPKKICLVGFCDEKEMPTQTLVPAVASKNVPALELVE 300
 QY 301 PIKKEVPAKSDYCEVCEFLVKEVTKLIDNNTEKEIILDAFPMKSKLPKLSSECOEY 360
 Db 301 PIKKEVPAKSDYCEVCEFLVKEVTKLIDNNTEKEIILDAFPMKSKLPKLSSECOEY 360
 QY 361 VDTYSSILSLLEEVSPPELVCSMLHLCSTRLPALTVHTQPKDGCFEVCCKLVGYLD 420
 Db 361 VDTYSSILSLLEEVSPPELVCSMLHLCSTRLPALTVHTQPKDGCFEVCCKLVGYLD 420
 QY 421 RNLKSTKQELILAELEKGSFLPDYQKQDOFVAEYEPVLEILVEWMDSPVCLTKIG 480
 Db 421 RNLKSTKQELILAELEKGSFLPDYQKQDOFVAEYEPVLEILVEWMDSPVCLTKIG 480
 QY 481 ACSNAKPLIGTEKCIWGSYWCQNTETAQCAVHCHCLOTWNNKPTVKS 524
 Db 481 ACSNAKPLIGTEKCIWGSYWCQNTETAQCAVHCHCLOTWNNKPTVKS 524

RESULT 5
 ABOUS200 standard; protein; 524 AA.
 ID ABOUS200 standard; protein; 524 AA.
 XX
 AC ABOUS200;
 XX
 DT 29-JAN-2003 (first entry)
 XX
 DE Human expressed protein tag (EPT) #1866.
 XX
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200278524-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 28-MAR-2002; 2002MO-US009671.
 XX
 PR 28-MAR-2001; 2001US-0279495P.
 XX
 PR 21-MAY-2001; 2001US-0292544P.
 XX
 PR 08-AUG-2001; 2001US-0310801P.
 XX
 PR 01-OCT-2001; 2001US-0326370P.
 XX
 PR 04-DEC-2001; 2001US-0336780P.
 XX
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 PA (ZYCO-) ZYCOs INC.
 XX
 PI Chicx RM, Tomlinson AJ, Urban RG;
 XX
 DR WPI; 2003-040607/03.
 XX

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.

XX Example 2; SEQ ID NO 1866; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 2,3e-237;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLASLIGALAGVILGLKECTRGSAVWCQNVKTASDCGAVKHLQTVWNNKPTVKS 60
DB 1 MYALFLASLIGALAGVILGLKECTRGSAVWCQNVKTASDCGAVKHLQTVWNNKPTVKS 60
QY 61 LPDCICKDVTAAGDMLKDNATEEELIYVLEKTCQMLPKPMNSASCKEIVSYLPVILI 120
DB 61 LPDCICKDVTAAGDMLKDNATEEELIYVLEKTCQMLPKPMNSASCKEIVSYLPVILI 120
QY 121 IKGEMSRGEVCSALNLCESLQKHLAEINHOKELESNKIPILDMTEVVAFPFMANIPLLLY 180
DB 121 IKGEMSRGEVCSALNLCESLQKHLAEINHOKELESNKIPILDMTEVVAFPFMANIPLLLY 180
QY 181 PODEGRSRPOKNDNDVQDCIQWVTDIQTAVRTNSTVQALVENYKCECRLEGMDI 240
DB 181 PODEGRSRPOKNDNDVQDCIQWVTDIQTAVRTNSTVQALVENYKCECRLEGMDI 240
QY 241 CKNYISOYSEIAIQMMHMOPEKICALVGFCEVEMPMQTLVPAKVASKVITPALLEVE 300
DB 241 CKNYISOYSEIAIQMMHMOPEKICALVGFCEVEMPMQTLVPAKVASKVITPALLEVE 300
QY 301 PIKKEHVPKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDMKCSKLPKSLSECEV 360
DB 301 PIKKEHVPKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDMKCSKLPKSLSECEV 360
QY 361 VDTYGSSTLILIEVSEBELVCSMTLCSGTRLPALTYHVTQPKGGGCEVCKLVGLD 420
DB 361 VDTYGSSTLILIEVSEBELVCSMTLCSGTRLPALTYHVTQPKGGGCEVCKLVGLD 420
QY 421 ENLEKNSRTOEILALEKGCSEFLPDYQKQCDQFAVEEVPVILITILIVEMDSFVCLKIG 480
DB 421 ENLEKNSRTOEILALEKGCSEFLPDYQKQCDQFAVEEVPVILITILIVEMDSFVCLKIG 480
QY 481 ACPSAHKPLIGTEKCIWGPSYWCQNTETPAQCNAAVEHCKRHVN 524
DB 481 ACPSAHKPLIGTEKCIWGPSYWCQNTETPAQCNAAVEHCKRHVN 524

RESULT 6
ABU05207
ID ABU05207 standard; protein; 524 AA.
XX
AC ABU05207;

XX 29-JAN-2003 (first entry)
DT
XX
XX Human expressed protein tag (EPT) #1873.
DE
XX
XX Translational profiling; expressed protein tag; EPT; Kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.

XX Homo sapiens.

XX WO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

XX 21-MAY-2001; 2001US-0292544P.

XX 08-AUG-2001; 2001US-0310801P.

XX 01-OCT-2001; 2001US-0326370P.

XX 04-DEC-2001; 2001US-0336780P.

XX 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCO INC.

XX Chicx RM, Tomlinson AV, Urban RG;

XX WPI, 2003-040607/03.

XX Example 2; SEQ ID NO 1873; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 2,3e-237;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLASLIGALAGVILGLKECTRGSAVWCQNVKTASDCGAVKHLQTVWNNKPTVKS 60
DB 1 MYALFLASLIGALAGVILGLKECTRGSAVWCQNVKTASDCGAVKHLQTVWNNKPTVKS 60
QY 61 LPDCICKDVTAAGDMLKDNATEEELIYVLEKTCQMLPKPMNSASCKEIVSYLPVILI 120
DB 61 LPDCICKDVTAAGDMLKDNATEEELIYVLEKTCQMLPKPMNSASCKEIVSYLPVILI 120
QY 121 IKGEMSRGEVCSALNLCESLQKHLAEINHOKELESNKIPILDMTEVVAFPFMANIPLLLY 180
DB 121 IKGEMSRGEVCSALNLCESLQKHLAEINHOKELESNKIPILDMTEVVAFPFMANIPLLLY 180

Db 121 IKGMSRPRGEVCSALNLCESIQKHLAEINHQLESNKIPEDMTETVAFPMANIPLLLY 180
Qy 181 PODPRSRKQPKNDGVCDQCIQAVTDIQTAVRTNSTFVQALVEHVEKECDRLGPGMADI 240
Db 181 PODPRSRKQPKNDGVCDQCIQAVTDIQTAVRTNSTFVQALVEHVEKECDRLGPGMADI 240
Qy 241 CKNYISQYSEIAIQMMHMQPKKEICALVGCDEVEKEMQOTLVPAKVASKNVIALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKKEICALVGCDEVEKEMQOTLVPAKVASKNVIALELVE 300
Qy 301 PIKHEVPAKSDVYCEVCEFLVEVTKLIDNNKTEKILDAFDKMSKLPKSLSECOEV 360
Db 301 PIKHEVPAKSDVYCEVCEFLVEVTKLIDNNKTEKILDAFDKMSKLPKSLSECOEV 360
Qy 361 VDTYGSIIILILEEVSPELVCSMLHLCSTRLPALTVHTQPKDGFCEVCKLVGYLD 420
Db 361 VDTYGSIIILILEEVSPELVCSMLHLCSTRLPALTVHTQPKDGFCEVCKLVGYLD 420
Qy 421 RNLEKSTKQELIALLEKGCFLPDYQKQDQFVAEYEPVLEILVEVMDPSFVCLKIG 480
Db 421 RNLEKSTKQELIALLEKGCFLPDYQKQDQFVAEYEPVLEILVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLGTGKCIWGPSYWCNTETAACNAVEHCKRHVWN 524
Db 481 ACPSAHKPLGTGKCIWGPSYWCNTETAACNAVEHCKRHVWN 524

RESULT 7
ABU05203
ID ABU05203 standard; protein; 524 AA.
XX
XX ABU05203;

DT 29-JAN-2003 (first entry)

DE Human expressed protein tag (EPT) #1869.

XX Translation profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.

OS Homo sapiens.

XX MO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

XX 21-MAY-2001; 2001US-0292544P.

XX 08-AUG-2001; 2001US-0310801P.

XX 01-OCT-2001; 2001US-0326370P.

XX 04-DEC-2001; 2001US-0336780P.

XX 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCO INC.

XX Chicz RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.

XX Example 2; SEQ ID NO 1869; 134p; English.

XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,

CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that bind to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polymulectins are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC ftp://ipo.int/pub/published_ppt_sequences

CC Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 2,3e-237;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALFLASLIGALAGPVLGLKECTRGSAWCONVKTASDCGAVHCLQTVWNRPTVKS 60
Db 1 MYALFLASLIGALAGPVLGLKECTRGSAWCONVKTASDCGAVHCLQTVWNRPTVKS 60

Qy 61 LPDICKDVVTAAGDMKNKATEEELVLYLEKTCMPLPKNNASCKEIVDSYLPIVLDI 120
Db 61 LPDICKDVVTAAGDMKNKATEEELVLYLEKTCMPLPKNNASCKEIVDSYLPIVLDI 120

Qy 121 IKGMSRPRGEVCSALNLCESIQKHLAEINHQLESNKIPEDMTETVAFPMANIPLLLY 180
Db 121 IKGMSRPRGEVCSALNLCESIQKHLAEINHQLESNKIPEDMTETVAFPMANIPLLLY 180

Qy 181 PODPRSRKQPKNDGVCDQCIQAVTDIQTAVRTNSTFVQALVEHVEKECDRLGPGMADI 240
Db 181 PODPRSRKQPKNDGVCDQCIQAVTDIQTAVRTNSTFVQALVEHVEKECDRLGPGMADI 240

Qy 241 CKNYISQYSEIAIQMMHMQPKKEICALVGCDEVEKEMQOTLVPAKVASKNVIALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKKEICALVGCDEVEKEMQOTLVPAKVASKNVIALELVE 300

Qy 301 PIKHEVPAKSDVYCEVCEFLVEVTKLIDNNKTEKILDAFDKMSKLPKSLSECOEV 360
Db 301 PIKHEVPAKSDVYCEVCEFLVEVTKLIDNNKTEKILDAFDKMSKLPKSLSECOEV 360

Qy 361 VDTYGSIIILILEEVSPELVCSMLHLCSTRLPALTVHTQPKDGFCEVCKLVGYLD 420
Db 361 VDTYGSIIILILEEVSPELVCSMLHLCSTRLPALTVHTQPKDGFCEVCKLVGYLD 420

Qy 421 RNLEKSTKQELIALLEKGCFLPDYQKQDQFVAEYEPVLEILVEVMDPSFVCLKIG 480
Db 421 RNLEKSTKQELIALLEKGCFLPDYQKQDQFVAEYEPVLEILVEVMDPSFVCLKIG 480

Qy 481 ACPSAHKPLGTGKCIWGPSYWCNTETAACNAVEHCKRHVWN 524
Db 481 ACPSAHKPLGTGKCIWGPSYWCNTETAACNAVEHCKRHVWN 524

RESULT 8
ABU07340
ID ABU07340 standard; protein; 524 AA.

XX ABU07340;

XX 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #2041.

XX Translation profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;

KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.
 XX
 OS Homo sapiens.
 XX
 PN MO200278524-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 28-MAR-2002; 2002WO-US009671.
 XX
 PR 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 PA (ZYCO-) ZYCOS INC.
 XX
 PI Chiciz RM, Tomlinson AJ, Urban RG;
 XX
 DR WPI; 2003-040607/03.
 XX
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX
 PS Example 2; SEQ ID NO 2041; 134pp; English.
 XX
 CC The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 524 AA:
 SQ
 Query Match 100.0%; Score 2789; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILLDFDKMCSKLPKSLSECOEV 360
 DB 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILLDFDKMCSKLPKSLSECOEV 360
 QY 361 VDTYSSILSLILEEVPPELVCSMLHCSGTRLPALTVHTVTPQKDGFCCEVCKLVGYLD 420
 DB 361 VDTYSSILSLILEEVPPELVCSMLHCSGTRLPALTVHTVTPQKDGFCCEVCKLVGYLD 420
 QY 421 RNLEKSTKOEILLALEKGSFLDPYQKQDGFVAEYEPVLIILVEMDPSFVCLKIG 480
 DB 421 RNLEKSTKOEILLALEKGSFLDPYQKQDGFVAEYEPVLIILVEMDPSFVCLKIG 480
 QY 481 ACPSAHKPLDGTCEKICWSPSYWCNTETAACNAVEHCKRHVN 524
 DB 481 ACPSAHKPLDGTCEKICWSPSYWCNTETAACNAVEHCKRHVN 524
 RESULT 9
 ID ABU05216 standard; protein; 524 AA.
 AC ABU05216;
 XX
 XX 29-JAN-2003 (first entry)
 DT
 XX Human expressed protein tag (EPT) #1882.
 DE
 XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.
 XX
 OS Homo sapiens.
 XX
 PN MO200278524-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 28-MAR-2002; 2002WO-US009671.
 XX
 PR 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 PA (ZYCO-) ZYCOS INC.
 XX
 PI Chiciz RM, Tomlinson AJ, Urban RG;
 XX
 DR WPI; 2003-040607/03.
 XX
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX
 PS Example 2; SEQ ID NO 1882; 134pp; English.
 XX
 CC The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,

CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: this sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MYALFLIASLIGALAGPYLGKECTRGSAVWCQNVKTAASDCGAVKCLQTVMNKPVTYS 60
DB 1 MYALFLIASLIGALAGPYLGKECTRGSAVWCQNVKTAASDCGAVKCLQTVMNKPVTYS 60
QY 61 LPEDICKDVVTAAGDMKDNATEEELIYVLEKTCMWLPRPMNSASCKEIVDSYLPVILDI 120
DB 61 LPEDICKDVVTAAGDMKDNATEEELIYVLEKTCMWLPRPMNSASCKEIVDSYLPVILDI 120
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAEINHOKELESNKIPBLDMTEVAPPMANIPILLY 180
DB 121 IKGEMSRPGEVCSALNLCESLQKHLAEINHOKELESNKIPBLDMTEVAPPMANIPILLY 180
QY 181 PODGPRSKPQPKNDGVCQDCIQWVTDIQTAVRNSFEVQALVHVKEECRLGPGMADI 240
DB 181 PODGPRSKPQPKNDGVCQDCIQWVTDIQTAVRNSFEVQALVHVKEECRLGPGMADI 240
QY 241 CKNYISQYSEIAIQMMHMOPEKICALVGFCEVKEPMQTLVPAKVASKNVITALELVE 300
DB 241 CKNYISQYSEIAIQMMHMOPEKICALVGFCEVKEPMQTLVPAKVASKNVITALELVE 300
QY 301 PIKKEHVPAKSDVYCEVCEFLVEKYVTKLIDNNKTEKEIIDAFDMKCSKLPSLSBECQEV 360
DB 301 PIKKEHVPAKSDVYCEVCEFLVEKYVTKLIDNNKTEKEIIDAFDMKCSKLPSLSBECQEV 360
QY 361 VDTYSSILSLILEEVSPELVCSMLHLCGSTRLPALTVHVTQPKDGFCEVCKLVGYLD 420
DB 361 VDTYSSILSLILEEVSPELVCSMLHLCGSTRLPALTVHVTQPKDGFCEVCKLVGYLD 420
QY 421 RNLEKNSTKOEIILAEKGCSTFLPPYQKOCQPFVAEVEPILILEVWMPDSFVCLKIG 480
DB 421 RNLEKNSTKOEIILAEKGCSTFLPPYQKOCQPFVAEVEPILILEVWMPDSFVCLKIG 480
QY 481 ACPSAHAKPLGTGECTIWGPSTWCQNTETAAQCNAVEHCKRHVMN 524
DB 481 ACPSAHAKPLGTGECTIWGPSTWCQNTETAAQCNAVEHCKRHVMN 524

```

RESULT 10
 ABU05202
 ID ABU05202 standard; protein, 524 AA.

XX ABU05202;
 XX 29-JAN-2003 (first entry)

DE Human expressed protein tag (EPT) #1868.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 XX protease; protease inhibitor; transporter; cytoskeletal protein;
 XX receptor; transcription factor; cancer; MHC;
 XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

OS Homo sapiens.
 XX PN WO200278524-A2.
 XX 10-OCT-2002.
 XX 28-MAR-2002; 2002MO-US009671.

XX 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCO INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukaemia.

XX Example 2; SEQ ID NO 1868; 134bp; English.

XX The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MYALFLIASLIGALAGPYLGKECTRGSAVWCQNVKTAASDCGAVKCLQTVMNKPVTYS 60
DB 1 MYALFLIASLIGALAGPYLGKECTRGSAVWCQNVKTAASDCGAVKCLQTVMNKPVTYS 60
QY 61 LPEDICKDVVTAAGDMKDNATEEELIYVLEKTCMWLPRPMNSASCKEIVDSYLPVILDI 120
DB 61 LPEDICKDVVTAAGDMKDNATEEELIYVLEKTCMWLPRPMNSASCKEIVDSYLPVILDI 120
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAEINHOKELESNKIPBLDMTEVAPPMANIPILLY 180
DB 121 IKGEMSRPGEVCSALNLCESLQKHLAEINHOKELESNKIPBLDMTEVAPPMANIPILLY 180
QY 181 PODGPRSKPQPKNDGVCQDCIQWVTDIQTAVRNSFEVQALVHVKEECRLGPGMADI 240
DB 181 PODGPRSKPQPKNDGVCQDCIQWVTDIQTAVRNSFEVQALVHVKEECRLGPGMADI 240
QY 241 CKNYISQYSEIAIQMMHMOPEKICALVGFCEVKEPMQTLVPAKVASKNVITALELVE 300
DB 241 CKNYISQYSEIAIQMMHMOPEKICALVGFCEVKEPMQTLVPAKVASKNVITALELVE 300
QY 301 PIKKEHVPAKSDVYCEVCEFLVEKYVTKLIDNNKTEKEIIDAFDMKCSKLPSLSBECQEV 360
DB 301 PIKKEHVPAKSDVYCEVCEFLVEKYVTKLIDNNKTEKEIIDAFDMKCSKLPSLSBECQEV 360
QY 361 VDTYSSILSLILEEVSPELVCSMLHLCGSTRLPALTVHVTQPKDGFCEVCKLVGYLD 420
DB 361 VDTYSSILSLILEEVSPELVCSMLHLCGSTRLPALTVHVTQPKDGFCEVCKLVGYLD 420

```

QY 421 RNLEKSTKQEIILALEKGCSEFLPDPYOKCQDFVAEYEPVLIIEIVMDSFVCLKIG 480
Db 421 RNLEKSTKQEIILALEKGCSEFLPDPYOKCQDFVAEYEPVLIIEIVMDSFVCLKIG 480
QY 481 ACPSAHKPELGTETKICWGPSYWCNTETPAQCNVAEHCKRHVMN 524
Db 481 ACPSAHKPELGTETKICWGPSYWCNTETPAQCNVAEHCKRHVMN 524

RESULT 11
ABU05214
ID ABU05214 standard; protein; 524 AA.
XX
AC ABU05214;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1880.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
FN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PE 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Chiciz RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
XX Example 2; SEQ ID NO 1880; 134pp; English.
PS
XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 2,3e-237;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLASLIGALAGVPLGLKECTRGSAVWCVNKTASDCGAVKHCLQTVMNKPITYKS 60
Db 1 MYALFLASLIGALAGVPLGLKECTRGSAVWCVNKTASDCGAVKHCLQTVMNKPITYKS 60
QY 61 LPDICI CKDVVTAAGDMLKDNATEEELIYLEKTCMDLKPKNMSASCKEIVDSYLPVILI 120
Db 61 LPDICI CKDVVTAAGDMLKDNATEEELIYLEKTCMDLKPKNMSASCKEIVDSYLPVILI 120
QY 121 IKGMSRPGEVCSALNICESELOKHLAEINHOKELENNKIPBLDMTEVVAAPPANIPILLY 180
Db 121 IKGMSRPGEVCSALNICESELOKHLAEINHOKELENNKIPBLDMTEVVAAPPANIPILLY 180
QY 181 PDGPRSKPQPKNDGVCQDCICQWVTDIQTAVRTNSTFVQALVBEHVECDRLGPGMADI 240
Db 181 PDGPRSKPQPKNDGVCQDCICQWVTDIQTAVRTNSTFVQALVBEHVECDRLGPGMADI 240
QY 241 CKNTISQYSEIATQMMHMQPEICALVGFCDVEKEMQTLVPAKVASKVIVIPALIVE 300
Db 241 CKNTISQYSEIATQMMHMQPEICALVGFCDVEKEMQTLVPAKVASKVIVIPALIVE 300
QY 301 PIKKHEVPAKSDVYCEVEPLVKEVTYKLDNNKTEKEILDAFDKVCSLPKSLSEECQEV 360
Db 301 PIKKHEVPAKSDVYCEVEPLVKEVTYKLDNNKTEKEILDAFDKVCSLPKSLSEECQEV 360
QY 361 VDTYGSSTLSLLEBSEVELVCSMLHLCSTGRLPALTVHTVQPKDGGCEVCCKLVGLD 420
Db 361 VDTYGSSTLSLLEBSEVELVCSMLHLCSTGRLPALTVHTVQPKDGGCEVCCKLVGLD 420
QY 421 RNLEKSTKQEIILALEKGCSEFLPDPYOKCQDFVAEYEPVLIIEIVMDSFVCLKIG 480
Db 421 RNLEKSTKQEIILALEKGCSEFLPDPYOKCQDFVAEYEPVLIIEIVMDSFVCLKIG 480
QY 481 ACPSAHKPELGTETKICWGPSYWCNTETPAQCNVAEHCKRHVMN 524
Db 481 ACPSAHKPELGTETKICWGPSYWCNTETPAQCNVAEHCKRHVMN 524

RESULT 12
ABU05215
ID ABU05215 standard; protein; 524 AA.
XX
AC ABU05215;
XX
XX
DT 29-JAN-2003 (first entry)
XX
XX Human expressed protein tag (EPT) #1881.
DE
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
FN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PE 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.

XX Chicz RM, Tomlinson AJ, Urban RG;
 XX WPI; 2003-040607/03.
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 XX leukemia.
 PS Example 2; SEQ ID NO 1881; 134pp; English.
 XX The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide, the purified polypeptide, or the antibody that binds to this
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptide is also
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 2,3e-237;
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLASLIGALAGVTLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKKPTVKS 60
 DB 1 MYALFLASLIGALAGVTLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKKPTVKS 60
 QY 61 LPDCICKDVTTAAGDMLKDNATEEIIIVYLEKTCQMLPKPMNSASCKEIVDSYLPVILI 120
 DB 61 LPDCICKDVTTAAGDMLKDNATEEIIIVYLEKTCQMLPKPMNSASCKEIVDSYLPVILI 120
 QY 121 IKGEMSRGEVCSALNLCESLQKHLAEINHQKOLESNKIPELDMTEVAVPMMANIPILLY 180
 DB 121 IKGEMSRGEVCSALNLCESLQKHLAEINHQKOLESNKIPELDMTEVAVPMMANIPILLY 180
 QY 181 FODGPRSRKPPKNDGVQDCICQWVTDIQTAVRNTSTFVQALVHVKEECDRLGCPMADI 240
 DB 181 FODGPRSRKPPKNDGVQDCICQWVTDIQTAVRNTSTFVQALVHVKEECDRLGCPMADI 240
 QY 241 CKNTIISQYSEIAIOMMMHMPKEICALVGFCEVKEPMQTLVPAKASKNVIPALERVE 300
 DB 241 CKNTIISQYSEIAIOMMMHMPKEICALVGFCEVKEPMQTLVPAKASKNVIPALERVE 300
 QY 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILLDAFDMCKSLPKXSISEECQEV 360
 DB 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILLDAFDMCKSLPKXSISEECQEV 360
 QY 361 VDIYSSILSLLEEVPELVCSMLHCSGTRLPALVHTVHTOPDDGFCFCVCKKLVGYLD 420
 DB 361 VDIYSSILSLLEEVPELVCSMLHCSGTRLPALVHTVHTOPDDGFCFCVCKKLVGYLD 420
 QY 421 RNLEKSTKQELLALAEKGSFLPDYQKODQFVAEYEPVLLIILVENVDPFVCLKTIG 480
 DB 421 RNLEKSTKQELLALAEKGSFLPDYQKODQFVAEYEPVLLIILVENVDPFVCLKTIG 480
 QY 481 ACSBAHKLGTETKICIMGPSYWCQNTETAACNAVHECKRHVNN 524
 DB 481 ACSBAHKLGTETKICIMGPSYWCQNTETAACNAVHECKRHVNN 524

RESULT 13
 AB05199
 ID AB05199 standard; protein, 524 AA.
 XX
 AC AB05199;
 XX
 DT 29-JAN-2003 (first entry)
 XX
 DE Human expressed protein tag (EPT) #1865.
 XX
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.
 OS Homo sapiens.
 XX
 PN WO200278524-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 28-MAR-2002; 2002WO-US009671.
 XX
 PR 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 PA (ZYCO-) ZYCOS INC.
 XX
 PI Chicz RM, Tomlinson AJ, Urban RG;
 XX
 DR WPI; 2003-040607/03.
 XX
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX
 PS Example 2; SEQ ID NO 1865; 134pp; English.
 XX The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide, the purified polypeptide, or the antibody that binds to this
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptide is also
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 2,3e-237;
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLASLIGALAGVTLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKKPTVKS 60
 DB 1 MYALFLASLIGALAGVTLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKKPTVKS 60
 QY 61 LPDCICKDVTTAAGDMLKDNATEEIIIVYLEKTCQMLPKPMNSASCKEIVDSYLPVILI 120


```
Db 61 LPCDICKDVTTAAGDMLKDNATEEELVLYLEKTCOMLEKPMNSASCKEIVSYLPVLLDI 120
Qy 121 IKGEMSRGGEVCSALNLCESIQKHIAELNHOKOLESNKIPELDMTEVAPPMANIPLLLY 180
Db 121 IKGEMSRGGEVCSALNLCESIQKHIAELNHOKOLESNKIPELDMTEVAPPMANIPLLLY 180
Qy 181 PDDGPRSKPQPKDNGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKEECRLGGMADI 240
Db 181 PDDGPRSKPQPKDNGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKEECRLGGMADI 240
Qy 241 CKNYISQYSEIAIQMMHMOPEKICALVGFCDVEKEMQTLVPAKASKNVVPALVELVE 300
Db 241 CKNYISQYSEIAIQMMHMOPEKICALVGFCDVEKEMQTLVPAKASKNVVPALVELVE 300
Qy 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMKCSKLPSLSSECOEV 360
Db 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMKCSKLPSLSSECOEV 360
Qy 361 VDTYSSILSLILEBVSPELVCSMLHLCSGTRLPALTVAHTVQPKDGGFCVCKLVGYLD 420
Db 361 VDTYSSILSLILEBVSPELVCSMLHLCSGTRLPALTVAHTVQPKDGGFCVCKLVGYLD 420
Qy 421 RNLEKNSTKOEILALAEKGCSEFLPDPYOKQCDQFVAEYEPVLIETLVEMDPSFYCLKIG 480
Db 421 RNLEKNSTKOEILALAEKGCSEFLPDPYOKQCDQFVAEYEPVLIETLVEMDPSFYCLKIG 480
Qy 481 ACPSAHKEPLGTERKCIWGPSYWCNTETFAACNAVEHCKRHVMN 524
Db 481 ACPSAHKEPLGTERKCIWGPSYWCNTETFAACNAVEHCKRHVMN 524
```

RESULT 14

```
ABU05212
ID ABU05212 standard; protein: 524 AA.
XX AC ABU05212;
XX DT 29-JAN-2003 (first entry)
XX DE Human expressed protein tag (EPT) #1878.
XX KM Translational profiling: expressed protein tag: EPT; kinase: phosphatase;
XX KM protease; protease inhibitor; transporter; cytoskeletal protein;
XX KM receptor; transcription factor; cancer; MHC;
XX KM major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX KM adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX OS Homo sapiens.
XX FN WO200278524-A2.
XX PD 10-OCT-2002.
XX PF 28-MAR-2002; 2002WO-US009671.
XX PR 28-MAR-2001; 2001US-0279495P.
XX PR 21-MAY-2001; 2001US-0292544P.
XX PR 08-AUG-2001; 2001US-0310801P.
XX PR 01-OCT-2001; 2001US-0326370P.
XX PR 04-DEC-2001; 2001US-0336780P.
XX PR 20-FEB-2002; 2002US-0358985P.
XX PA (ZYCO-) ZYCOs INC.
XX PI Chiciz RM, Tomlinson AJ, Urban RG;
XX DR WPI; 2003-040607/03.
XX DE New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
```

```
XX PS Example 2; SEQ ID NO 1878; 134pp; English.
XX CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 524 AA;
```

Query Match 100.0%; Score 2789; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 2,3e-27;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MYALFLASLIGALAGVGLGKECTRGSAVWQNVKTASDCGAVKCLQTVWNNPTVKS 60
Db 1 MYALFLASLIGALAGVGLGKECTRGSAVWQNVKTASDCGAVKCLQTVWNNPTVKS 60
Qy 61 LPCDICKDVTTAAGDMLKDNATEEELVLYLEKTCOMLEKPMNSASCKEIVSYLPVLLDI 120
Db 61 LPCDICKDVTTAAGDMLKDNATEEELVLYLEKTCOMLEKPMNSASCKEIVSYLPVLLDI 120
Qy 121 IKGEMSRGGEVCSALNLCESIQKHIAELNHOKOLESNKIPELDMTEVAPPMANIPLLLY 180
Db 121 IKGEMSRGGEVCSALNLCESIQKHIAELNHOKOLESNKIPELDMTEVAPPMANIPLLLY 180
Qy 181 PDDGPRSKPQPKDNGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKEECRLGGMADI 240
Db 181 PDDGPRSKPQPKDNGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKEECRLGGMADI 240
Qy 241 CKNYISQYSEIAIQMMHMOPEKICALVGFCDVEKEMQTLVPAKASKNVVPALVELVE 300
Db 241 CKNYISQYSEIAIQMMHMOPEKICALVGFCDVEKEMQTLVPAKASKNVVPALVELVE 300
Qy 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMKCSKLPSLSSECOEV 360
Db 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMKCSKLPSLSSECOEV 360
Qy 361 VDTYSSILSLILEBVSPELVCSMLHLCSGTRLPALTVAHTVQPKDGGFCVCKLVGYLD 420
Db 361 VDTYSSILSLILEBVSPELVCSMLHLCSGTRLPALTVAHTVQPKDGGFCVCKLVGYLD 420
Qy 421 RNLEKNSTKOEILALAEKGCSEFLPDPYOKQCDQFVAEYEPVLIETLVEMDPSFYCLKIG 480
Db 421 RNLEKNSTKOEILALAEKGCSEFLPDPYOKQCDQFVAEYEPVLIETLVEMDPSFYCLKIG 480
Qy 481 ACPSAHKEPLGTERKCIWGPSYWCNTETFAACNAVEHCKRHVMN 524
Db 481 ACPSAHKEPLGTERKCIWGPSYWCNTETFAACNAVEHCKRHVMN 524
```

RESULT 15

```
ABU05213
ID ABU05213 standard; protein: 524 AA.
XX AC ABU05213;
XX DT 29-JAN-2003 (first entry)
XX DE Human expressed protein tag (EPT) #1879.
```

XX Translational profiling; expressed protein tag; EPT, kinase; phosphatase;
 KM protease; protease inhibitor; transporter; cytoskeletal protein;
 KM receptor; transcription factor; cancer; MHC;
 KM major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KM adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia;
 OS Homo sapiens.
 XX MO200278524-A2.
 XX 10-OCT-2002.
 XX 28-MAR-2002; 2002WO-US009671.
 XX 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX (ZYCO-) ZYCO INC.
 PA Chicx RM, Tomlinson AJ, Urban RG;
 PI WPI; 2003-040607/03.
 DR New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 XX cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX Example 2; SEQ ID NO 1879; 134pp; English.
 PS The invention describes a purified polypeptide, which comprises a
 XX fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 524 AA;
 SQ
 Query Match 100.0%; Score 2789; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MYALFLLASLLGALAGPVLGLKECTGSAWCONVKTASDCGAKHCLQTVNNKPTVKS 60
 Db 1 MYALFLLASLLGALAGPVLGLKECTGSAWCONVKTASDCGAKHCLQTVNNKPTVKS 60
 QY 61 LPDCICKDVYTAAGDMKDNANTEBEILLVLEKTCDMILPKPNMSASCKEIVDSYLPVLDI 120
 Db 61 LPDCICKDVYTAAGDMKDNANTEBEILLVLEKTCDMILPKPNMSASCKEIVDSYLPVLDI 120
 QY 121 IKGEMSRPGEVCSALNCEISLOKHLAEINOKOLESNKIPELDMTEVAPFMANIPILLY 180
 Db 121 IKGEMSRPGEVCSALNCEISLOKHLAEINOKOLESNKIPELDMTEVAPFMANIPILLY 180
 QY 181 PDGPRSKPQPKNGDVCCODCIQWVTDIQTAVRTNSTFVQALVEHVEECDRILGPGMADI 240
 Db 181 PDGPRSKPQPKNGDVCCODCIQWVTDIQTAVRTNSTFVQALVEHVEECDRILGPGMADI 240

Db 181 PDGPRSKPQPKNGDVCCODCIQWVTDIQTAVRTNSTFVQALVEHVEECDRILGPGMADI 240
 QY 241 CKNYISOYSEIAIQMMHMPKEICALVGFCDVKEKMPQTLVPAKASKNVLPALVELVE 300
 Db 241 CKNYISOYSEIAIQMMHMPKEICALVGFCDVKEKMPQTLVPAKASKNVLPALVELVE 300
 QY 301 PIKHEVPASDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMCSKLPKSLSEECOEY 360
 Db 301 PIKHEVPASDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMCSKLPKSLSEECOEY 360
 QY 361 VDTGSSILSTLLEVSPELVCSMLHCSGTRLPALVHTQPKDGFCEVCCKLVGYLD 420
 Db 361 VDTGSSILSTLLEVSPELVCSMLHCSGTRLPALVHTQPKDGFCEVCCKLVGYLD 420
 QY 421 RNLEKSTKQEIILAEKGSFLPDYQKQDQFVAEYEPVLJEILVEWMDPSFVCLKIG 480
 Db 421 RNLEKSTKQEIILAEKGSFLPDYQKQDQFVAEYEPVLJEILVEWMDPSFVCLKIG 480
 QY 481 ACPSAHPPLIGTEKCIWGPSTWCONTEETAACNAVEHCKRHVNN 524
 Db 481 ACPSAHPPLIGTEKCIWGPSTWCONTEETAACNAVEHCKRHVNN 524

Search completed: May 5, 2004, 13:31:03
 Job time : 75.151 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:41:29 ; Search time 38.2188 Seconds
(without alignments)
4325.921 Million cell updates/sec

Title: US-09-743-684A-1
Perfect score: 524
Sequence: 1 MYALFLSLGALAGPVL.....NTETAACNAVEHCKRHVN 524

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL 25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	4-2	554	11 Q8BFQ1	Q8bfq1 mus musculus
2	20	3-8	121	6 P79254	P79254 ovis aries
3	13	2-5	13	4 O75905	O75905 homo sapien
4	13	2-5	512	13 Q7S70	Q7sy70 xenopus lae
5	12	2-3	449	11 Q8BJV5	Q8bjv5 mus musculus
6	12	2-3	525	11 Q8C1C1	Q8c1c1 mus musculus
7	11	2-1	402	11 Q8C1N0	Q8c1n0 mus musculus
8	9	1-7	520	13 Q8UWZ4	Q8uwz4 brachydanio
9	9	1-7	522	13 Q9DG82	Q9dgs2 brachydanio
10	8	1-5	50	9 Q854G1	Q854g1 mycobacteri
11	8	1-5	90	9 Q9ZXL7	Q9zxl7 bacterioph
12	8	1-5	154	9 O21999	O21999 bacterioph
13	8	1-5	162	16 Q928C7	Q928c7 listeria in
14	8	1-5	165	16 Q986B3	Q986b3 rhizobium l
15	8	1-5	177	4 Q9NSM6	Q9nsm6 homo sapien
16	8	1-5	179	2 Q47224	Q47224 escherichia

Q8cvg2 escherichia
Q82w66 nitrosomona
Q84gh7 gamma-prote
Q44889 bordetella
Q7wgys bordetella
Q7uaq7 shigella fl
Q8xc62 escherichia
Q83p69 shigella fl
Q91yq4 arabidopsis
Q48998 oryza sativ
Q88si0 brachydanio
Q7vpns haemophilus
Q93un6 vibrio vuln
Q8d492 vibrio vuln
Q8x584 escherichia
Q83rp7 shigella fl
Q9x7g2 methylobact
Q8jty7 lumpy skin
Q8jtn0 lumpy skin
Q91ml5 lumpy skin
Q8ybg5 brucella me
Q937r1 brucella me
Q8fwm3 brucella su
Q81243 salmonella
Q84922 escherichia
Q7wti0 escherichia
Q9exn9 escherichia
Q962k0 plasmodium
Q9f170 salmonella

ALIGNMENTS

RESULT 1

Q8BFQ1 ID Q8BFQ1 PRELIMINARY; PRT; 554 AA.
AC Q8BFQ1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Prosaposin.
GN PSAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Kidney, and Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
RL 60,770 full-length cDNAs."
DR EMBL; AK088369; BAC40308.1; -
DR MGD; MGI:97783; Psap.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; SapA.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR Pfam; PF02199; SApA; 2.
DR Pfam; PF05184; SapB_1; 4.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SApA; 2.

```

DR SMART; SM00118; SAPB; 4.
SQ SEQUENCE 554 AA; 61050 MW; FF58DB79C7CC0C18 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 1e-12; Length 554;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 ETLAALKGCSELPDPYQKQCD 452
Db 461 ETLAALKGCSELPDPYQKQCD 482

RESULT 2
P79254
ID P79254 PRELIMINARY; PRT; 121 AA.
AC P79254;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DE Sulfated glycoprotein-1/SGP-1 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96249303; PubMed=8848570;
RA Spencer T.E., Graf G.H., Bazer F.W.;
RT "Sulfated glycoprotein-1 (SGP-1) expression in ovine endometrium
during the oestrous cycle and early pregnancy.";
RL Reprod. Fert. Dev. 7:1053-1060(1995).
DR EMBL; S82555; AAD14405.1; -.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008373; Saposin.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR01797; SAPOSIN.
DR SMART; SM00118; SAPB; 1.
FT NON TER 1
SQ SEQUENCE 121 AA; 13604 MW; 4F0F5A6EB83D0C9A CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 2.7e-11; Length 121;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 HVKECDRLGPGMADICKNY 244
Db 21 HVKECDRLGPGMADICKNY 40

RESULT 3
O75905
ID O75905 PRELIMINARY; PRT; 13 AA.
AC O75905; O9HAQ9;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Prosaposin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98426155; PubMed=9751800;
RA Sun Y., Jin P., Witte D.P., Grabowski G.A.;
RT "Isolation and characterization of the human prosaposin promoter.";
RL Gene 218:37-47(1998).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RA Sun Y., Jin P., Grabowski G.A.;
RT "Isolation and characterization of the human prosaposin promoter.";
RL Gene 0:0-0(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21205763; PubMed=11309366;
RA Hulkova H., Cervenkova M., Ledvinov J., Tochackova M., Hrebicek M.,
RA Poupetova H., Befekadu A., Berna L., Paton B.C., Harzer K., Boeoz A.,
RA Smid P., Ellender M.;
RT "A novel mutation in the coding region of the prosaposin gene leads to
a complete deficiency of prosaposin and saposins, and is associated
with a complex sphingolipidosis dominated by lactosylceramide
accumulation.";
RL Hum. Mol. Genet. 10:927-940(2001).
DR EMBL; AF057307; AAC64921.2; -.
DR EMBL; AF307850; AAC31635.1; -.
FT NON TER 13
SQ SEQUENCE 13 AA; 1383 MW; 57585B6C1784D727 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 3.6e-05; Length 13;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLLASLLGA 13
Db 1 MYALFLLASLLGA 13

RESULT 4
O7SV70
ID O7SV70 PRELIMINARY; PRT; 512 AA.
AC O7SV70;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting R., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Whole;

RA Klein S., Strausberg R., to the EMBL/GenBank/DBJ databases.

RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC054988; AAH54988.1; -.

KW Hypothetical protein.

SQ SEQUENCE 512 AA; 57555 MW; 57CFASE2093F6FB1 CRC64;

Query Match

Best Local Similarity 2.5%; Score 13; DB 13; Length 512;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 52 VWNKPTVKSLPCD 64

Db 52 VWNKPTVKSLPCD 64

RESULT 5

Q8BJV5

ID Q8BJV5 PRELIMINARY; PRT; 449 AA.

AC Q8BJV5;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical saposin A-type domain/saposin type B containing

DE protein.

GN 2310020A21RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Eye;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573 (2002).

DR EMBL; AK078699; BAC37363.1; -.

DR MGD; MGI:1924193; 2310020A21RIK.

DR GO; GO:0005764; C:lysosome; IEA.

DR GO; GO:0006665; P:sphingolipid metabolism; IEA.

DR InterPro; IPR007856; SapB_1.

DR InterPro; IPR007856; SapB_2.

DR InterPro; IPR008138; SapB_sub.

DR InterPro; IPR008373; Saposin.

DR Pfam; PF02199; SAPA; 1.

DR Pfam; PF05184; SapB_1; 3.

DR Pfam; PF03489; SapB_2; 3.

DR PRINTS; PR01797; SAPOSIN.

DR ProDom; PD001732; SapB_sub; 2.

DR SMART; SM00162; SAPA; 1.

DR SMART; SM00118; SAPP; 3.

KW Hypothetical protein.

SQ SEQUENCE 449 AA; 48705 MW; 377CC42A475B292F CRC64;

Query Match

Best Local Similarity 2.3%; Score 12; DB 11; Length 449;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 52 VWNKPTVKSLPC 63

Db 56 VWNKPTVKSLPC 67

RESULT 6

Q8C1C1

ID Q8C1C1 PRELIMINARY; PRT; 525 AA.

AC Q8C1C1;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical saposin A-type domain/saposin type B containing

DE protein.

GN 2310020A21RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Tongue;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573 (2002).


```
Query Match 1.5%; Score 8; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 ILDIKGE 124
Db 43 ILDIKGE 50
|||||

RESULT 11
Q9ZXL7 PRELIMINARY; PRT; 90 AA.
AC Q9ZXL7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Orl0.
OS Bacteriophage phi CTX.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OC NCBI_TaxID=35343;
RN [1]
RP STRAIN=phiCTX-c;
RC Hayashi T.;
RA Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=phiCTX-c;
RX MEDLINE=92157549; PubMed=10027959;
RA Nakayama K., Kanaya S., Ohnishi M., Terawaki Y., Hayashi T.;
RT "The complete nucleotide sequence of phiCTX, a cytotoxin-converting
RT phage of Pseudomonas aeruginosa: implications for phage evolution and
RT horizontal gene transfer via bacteriophage.";
RL Mol. Microbiol. 31:399-419(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=phiCTX-c;
RX MEDLINE=90014160; PubMed=2507866;
RA Hayashi T., Kamio Y., Hishinuma F., Usami Y., Titani K., Terawaki Y.;
RT "Pseudomonas aeruginosa cytotoxin: the nucleotide sequence of the gene
RT and the mechanism of activation of the protoxin.";
RL Mol. Microbiol. 3:861-868(1989).
DR EMBL; AB008550; BAA36236.1; -
DR InterPro; IPR008473; DUF754.
DR Pfam; PF05449; DUF754; 1.
SQ SEQUENCE 90 AA; 9698 MW; 7E8C73C84E24F4D4 CRC64;

Query Match 1.5%; Score 8; DB 9; Length 90;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LLASLLGA 13
Db 37 LLASLLGA 44
|||||

RESULT 12
O21999 PRELIMINARY; PRT; 154 AA.
AC O21999;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE LacZ-alpha.
OS Bacteriophage lambda.
OG Plasmid pAL-Z.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OC NCBI_TaxID=10710;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=IS1;
RX MEDLINE=97473533; PubMed=9332387;
```

```
RA Ahmed A., Podemski L.;
RT "Use of ordered deletions in genome sequencing.";
RL Gene 197:367-373(1997).
DR EMBL; U66885; AAC48874.1; -
DR HSP; P00722; IBGL.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl...; IEA.
DR GO; GO:0015520; F:tetracycline:hydrogen antiporter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0015904; P:tetracycline transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR008979; Gal bind like.
DR InterPro; IPR006104; Glyco hydro 2SB.
DR InterPro; IPR005829; Sug transporter.
DR InterPro; IPR001958; TCR_TetA.
DR Pfam; PF02837; Glyco hydro_2_N; 1.
DR PRINTS; PR01035; TCRTETA.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW Plasmid.
SQ SEQUENCE 154 AA; 17242 MW; 2CEE57E8FA128370 CRC64;

Query Match 1.5%; Score 8; DB 9; Length 154;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LLASLLGA 13
Db 123 LLASLLGA 130
|||||

RESULT 13
Q928C7 PRELIMINARY; PRT; 162 AA.
AC Q928C7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein lin2609.
GN Lin2609.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OC NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Etian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596173; CAC97836.1; -
DR PIR; AD1758; AD1758.
DR Listlist; LIN02609; -
DR InterPro; IPR000437; Prok lipoprot S.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 162 AA; 18193 MW; 181DA681AE002BC7 CRC64;

Query Match 1.5%; Score 8; DB 16; Length 162;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 4 LFLASLL 11
|||||
Db 4 LFLASLL 11

Db 146 EILAALEK 153

Search completed: May 5, 2004, 13:47:43
Job time : 39.2188 secs

RESULT 14

Q986B3 ID Q986B3 PRELIMINARY; PRT; 165 AA.
AC Q986B3;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein mll7437.
GN MLL7437.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Ideawara K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003011; BAB53540.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 165 AA; 18468 MW; B8FA62EEAA271361 CRC64;

Query Match 1.5%; Score 8; DB 16; Length 165;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LLASLLGA 13
|||||
Db 49 LLASLLGA 56

RESULT 15

Q9NSM6 ID Q9NSM6 PRELIMINARY; PRT; 177 AA.
AC Q9NSM6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP761F1921.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amalgam;
RA Poustka A., Wellenreuther R., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; ALJ162013; CAB82365.1; -.
DR PIR; T47133; T47133.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 177 AA; 20522 MW; EB717DE9993AF417 CRC64;

Query Match 1.5%; Score 8; DB 4; Length 177;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 EILAALEK 438
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:35:24 ; Search time 10.5605 Seconds
(without alignments)
2583.672 Million cell updates/sec

Title: US-09-743-684A-1

Perfect score: 524

Sequence: 1 MYALFLIALSLGALAGVL.....NTETAQCNAVEHCKRHVMN 524

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	524	100.0	524	1	SAP HUMAN
2	30	5.7	525	1	SAP_BOVIN
3	22	4.2	80	1	SAP_PIG
4	22	4.2	554	1	SAP RAT
5	22	4.2	557	1	SAP_MOUSE
6	12	2.3	518	1	SAP_CHICK
7	9	1.7	81	1	SAP_CAVPO
8	8	1.5	179	1	FIMI_ECOLI
9	8	1.5	186	1	HS23_DROME
10	8	1.5	201	1	FMFX_BORPE
11	8	1.5	216	1	Y726_METJA
12	8	1.5	349	1	BOR4_MOUSE
13	8	1.5	356	1	BOR4_HUMAN
14	8	1.5	396	1	TCR3_ECOLI
15	8	1.5	435	1	MNTC_BACSU
16	8	1.5	490	1	FA10_RABIT
17	8	1.5	655	1	YHCP_ECOLI
18	8	1.5	813	1	FCR2_XENLA
19	7	1.3	84	1	ACPX_STRAW
20	7	1.3	105	1	RN37_YEAST
21	7	1.3	112	1	M1P_CHICK
22	7	1.3	137	1	CC23_HOMAM
23	7	1.3	144	1	Y638_TREPA
24	7	1.3	145	1	PTMA_BACST
25	7	1.3	158	1	ILVH_BUCAI
26	7	1.3	159	1	ILVH_BUCBP
27	7	1.3	164	1	ID1_RAT
28	7	1.3	168	1	ID1_MOUSE
29	7	1.3	174	1	NU6M_HUMAN
30	7	1.3	203	1	SODM_THETH
31	7	1.3	208	1	MOBA_RHILIO
32	7	1.3	210	1	AFS1_SCHPO
33	7	1.3	210	1	YB98_SCHPO

RESULT 1
SAP_HUMAN

ID	SAP_HUMAN	STANDARD;	PRT;	524 AA.
AC	P07602; P07292; P15793; P78538; P78541; P78546; P78547; P78558;			
AC	Q92739; Q92740; Q92741; Q92742;			
DT	01-APR-1988 (Rel. 07, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Proactivator polypeptide precursor [Contains: Saposin A (Protein A); Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside sulfate activator) (CSact) (Dispersin) (Sulfatide/GMI activator); Saposin C (Co-beta-glucosidase) (Al activator) (Glucosylceramidase activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D (Protein C) (Component C)].			
DE	PSAP.			
GN	Homo sapiens (Human).			
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=90129043; PubMed=2515150;			
RA	Rorman E.G., Grabowski G.A.;			
RT	"Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphingolipid hydrolase activator proteins are encoded by single genes in humans and rats.";			
RL	Genomics 5:486-492(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89255151; PubMed=2498298;			
RA	Nakano T., Sandhoff K., Stuenkel J., Christomanou H., Suzuki K.;			
RT	"Structure of full-length cDNA coding for sulfatide activator, a Co-beta-glucosidase and two other homologous proteins: two alternate forms of the sulfatide activator.";			
RL	J. Biochem. 105:152-154(1989).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM SAP-MU-0).			
RC	TISSUE=Brain, Eye, and Skin;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K., Altschul S.F., Zeeberg B., Buetow K.H., Moore T., Max S.I., Hsieh F., Hopkins R.F., Jordan H., Moore T., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny K.C., Hale S., Sodergren E.J., Lu X., Gibbs R.A., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,			

P74368 synechocyst
P23958 phasolus v
P42270 escherichia
P15764 cyanophora
P77373 escherichia
Q9hxq6 pseudomonas
Q88mvo pseudomonas
P75485 mycoplasma
Q9hvv5 pseudomonas
P31118 paramecium
Q9juu2 neisseria m
Q9jz80 neisseria m

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [4]
 RP SEQUENCE OF 59-125 AND 304-513 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=91192146; PubMed=2013321;
 RA Holschmidt H., Sandhoff K., Fuerst W., Kwon H.Y., Schnabel D.,
 RA Suzuki K.;
 RT "The organization of the gene for the human cerebroside sulfate
 activator protein.";
 RL FEBS Lett. 280:267-270 (1991).
 RN [5]
 RP SEQUENCE OF 164-524 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=88068647; PubMed=2825202;
 RA Dewji N.N., Wenger D.A., O'Brien J.S.;
 RT "Nucleotide sequence of cloned cDNA for human sphingolipid activator
 protein 1 precursor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8652-8656 (1987).
 RN [6]
 RP PARTIAL SEQUENCE OF 60-142.
 RX MEDLINE=89240739; PubMed=2717620;
 RA Morimoto S., Martin B.M., Yamamoto Y., Kretz K.A., O'Brien J.S.,
 RA Kishimoto Y.;
 RT "Saposin A: second cerebroside activator protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3389-3393 (1989).
 RN [7]
 RP SEQUENCE OF 195-263 FROM N.A.
 RX MEDLINE=86130593; PubMed=2869718;
 RA Dewji N.N., Wenger D.A., Fujibayashi S., Donoviel M., Esch F.,
 RA Hill F., O'Brien J.S.;
 RT "Molecular cloning of the sphingolipid activator protein-1 (SAP-1),
 the sulfate sulfate activator.";
 RL Biochem. Biophys. Res. Commun. 134:989-994 (1986).
 RN [8]
 RP SEQUENCE OF 195-274.
 RC TISSUE=Kidney;
 RX MEDLINE=91006165; PubMed=2209618;
 RA Furst W., Schubert J., Machleidt W., Meyer H.E., Sandhoff K.;
 RT "The complete amino-acid sequences of human ganglioside GM2 activator
 protein and cerebroside sulfate activator protein.";
 RL Eur. J. Biochem. 192:709-714 (1990).
 RN [9]
 RP SEQUENCE OF 195-274.
 RX MEDLINE=89207118; PubMed=3242555;
 RA Kleinschmidt T., Christomanou H., Braunitzer G.;
 RT "Complete amino-acid sequence of the naturally occurring A2 activator
 protein for enzymic sphingomyelin degradation: identity to the
 sulfate activator protein (SAP-1).";
 RL Biol. Chem. Hoppe-Seyler 369:1361-1365 (1988).
 RN [10]
 RP SEQUENCE OF 311-390.
 RX MEDLINE=88163077; PubMed=3442600;
 RA Kleinschmidt T., Christomanou H., Braunitzer G.;
 RT "Complete amino-acid sequence and carbohydrate content of the
 naturally occurring glucosylceramide activator protein (A1 activator)
 absent from a new human Gaucher disease variant.";
 RL Biol. Chem. Hoppe-Seyler 368:1571-1578 (1987).
 RN [11]
 RP SEQUENCE OF 407-484.
 RX MEDLINE=89000190; PubMed=3048308;
 RA Furst W., Machleidt W., Sandhoff K.;
 RT "The precursor of sulfate activator protein is processed to three
 different proteins.";
 RL Biol. Chem. Hoppe-Seyler 369:317-328 (1988).
 RN [12]
 RP PARTIAL SEQUENCE OF 405-484.
 RX MEDLINE=89025876; PubMed=2845979;
 RA Morimoto S., Martin B.M., Kishimoto Y., O'Brien J.S.;
 RT "Saposin D: a sphingomyelinase activator.";
 RL Biochem. Biophys. Res. Commun. 156:403-410 (1988).
 RN [13]

RP SEQUENCE OF 17-26.
 RC TISSUE=Milk;
 RX MEDLINE=92068206; PubMed=1958198;
 RA Kondo K., Hineno T., Sano A., Kakimoto Y.;
 RT "Isolation and characterization of prosaposin from human milk.";
 RL Biochem. Biophys. Res. Commun. 181:286-292 (1991).
 RN [14]
 RP PARTIAL SEQUENCE (SAPOSIN B), AND STRUCTURE OF CARBOHYDRATES.
 RC TISSUE=Urine;
 RX MEDLINE=20032116; PubMed=10562467;
 RA Fluhrty A.L., Lombardo C., Louis A., Stevens R.L., Whitelegge J.P.,
 RA Waring A.J., To T., Fluhrty C.B., Faulk K.F.;
 RT "Preparation of the cerebroside sulfate activator (CSAct or saposin B)
 from human urine.";
 RL Mol. Genet. Metab. 68:391-403 (1999).
 RN [15]
 RP STRUCTURE OF CARBOHYDRATE ON ASN-215.
 RX MEDLINE=21110404; PubMed=11180632;
 RA Faulk K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,
 RA Stevens R.L., Fluhrty C.B., Fluhrty A.L.;
 RT "Structure of the asparagine-linked sugar chains of porcine kidney and
 human urine cerebroside sulfate activator protein.";
 RL J. Mass Spectrom. 35:1416-1424 (2000).
 RN [16]
 RP SAPOSIN D DISULFIDE BONDS.
 RX MEDLINE=99337688; PubMed=10406958;
 RA Tatti M., Salvioli R., Ciaffoni F., Pucci P., Andolfo A.,
 RA Amoresano A., Vaccaro A.M.;
 RT "Structural and membrane-binding properties of saposin D.";
 RL Eur. J. Biochem. 263:486-494 (1999).
 RN [17]
 RP SAPOSIN B DISULFIDE BONDS.
 RX MEDLINE=22398398; PubMed=12510003;
 RA Ahn V.E., Faulk K.F., Whitelegge J.P., Higginson J., Fluhrty A.L.,
 RA Prive G.G.;
 RT "Expression, purification, crystallization, and preliminary X-ray
 analysis of recombinant human saposin B.";
 RL Protein Expr. Purif. 27:186-193 (2003).
 RN [18]
 RP MASS SPECTROMETRY.
 RC TISSUE=Urine;
 RX MEDLINE=99441404; PubMed=10510427;
 RA Faulk K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,
 RA Krutchinsky A.N., Standing K.G., Waring A.J., Stevens R.L.,
 RA Fluhrty C.B., Fluhrty A.L.;
 RT "Cerebroside sulfate activator protein (Saposin B): chromatographic
 and electrospray mass spectrometric properties.";
 RL J. Mass Spectrom. 34:1040-1054 (1999).
 RN [19]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 195-273, AND MUTAGENESIS OF
 ILE-240.
 RX MEDLINE=22406333; PubMed=12518053;
 RA Ahn V.E., Faulk K.F., Whitelegge J.P., Fluhrty A.L., Prive G.G.;
 RT "Crystal structure of saposin B reveals a dimeric shell for lipid
 binding.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:38-43 (2003).
 RN [20]
 RP REVIEW ON MLD VARIANTS.
 RX MEDLINE=95170731; PubMed=7866401;
 RA Gieselmann V., Zlotogora J., Harris A., Wenger D.A., Morris C.P.;
 RT "Molecular genetics of metachromatic leukodystrophy.";
 RL Hum. Mutat. 4:233-242 (1994).
 RN [21]
 RP VARIANT MLD ILE-217.
 RX MEDLINE=90147748; PubMed=2302219;
 RA Rafi M.A., Zhang X.-L., Degala G., Wenger D.A.;
 RT "Detection of a point mutation in sphingolipid activator protein-1
 mRNA in patients with a variant form of metachromatic
 leukodystrophy.";
 RL Biochem. Biophys. Res. Commun. 166:1017-1023 (1990).
 RN [22]
 RP SEQUENCE FROM N.A., AND VARIANT MLD ILE-217.
 RX MEDLINE=90207231; PubMed=2320574;

```
Query Match      100.0%; Score 524; DB 1; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALFLASLLGAALAGPVLGKECTRGSAVWCQVKTASDCGAVKHCLOQTWNKPTVKS 60
Db 1 MYALFLASLLGAALAGPVLGKECTRGSAVWCQVKTASDCGAVKHCLOQTWNKPTVKS 60
Qy 61 LPDCICKDVVTAAGDMLKDNATEEILVYLEKTCMDLPPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPDCICKDVVTAAGDMLKDNATEEILVYLEKTCMDLPPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLY 180
Qy 181 PDGPRSKPQKONGDVCDQCQDIQMTVDIOTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Db 181 PDGPRSKPQKONGDVCDQCQDIQMTVDIOTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Qy 241 CKNYISOYSEIAIQMMHMQPKEICALVGFCDVEKEMPOTLVPAKVASKNVPALVELVE 300
Db 241 CKNYISOYSEIAIQMMHMQPKEICALVGFCDVEKEMPOTLVPAKVASKNVPALVELVE 300
Qy 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPSLSEECQEV 360
Db 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPSLSEECQEV 360
Qy 361 VDTYSSILSILLBEVSPELVCSMLHLCSTGTRLPALTTHVTPKDGGFCEVCKKLGYLD 420
Db 361 VDTYSSILSILLBEVSPELVCSMLHLCSTGTRLPALTTHVTPKDGGFCEVCKKLGYLD 420
Qy 421 RNLKNTKTEILAALEKSGSFLPDYKQKCDQFVAEYEPVLIILVEVMDPSVCLKIG 480
Db 421 RNLKNTKTEILAALEKSGSFLPDYKQKCDQFVAEYEPVLIILVEVMDPSVCLKIG 480
Qy 481 ACPSAHPKLLGTGKXCINGPSYWCQNTETAACQNAVEHCKRHWVN 524
Db 481 ACPSAHPKLLGTGKXCINGPSYWCQNTETAACQNAVEHCKRHWVN 524

RESULT 2
SAP_BOVIN
ID - SAP BOVIN STANDARD; PRT; 525 AA.
AC P26779; Q9N2G4;
DT 01-AUG-1992 (Rel. 23, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Proactivator polypeptide precursor [Contains: Saposin A (Protein A);
DE Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside
DE sulfate activator) (CSact) (Dispersin) (Sulfatide/GMI activator);
DE Saposin C (Co-beta-glucosidase) (Al activator) (Glucosylceramidase
DE activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D
DE (Protein C) (Component C)].
GN PSAP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS.
RC TISSUR=Mammary gland;
RA Azuma N., Yoshida K.;
RT "RT-PCR cloning of bovine proasaposin.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 312-391.
RC TISSUR=Spleen;
RA MEDLINE=92207994; PubMed=1554743;
RA Sano A., Mizuno T., Kondoh K., Hineno T., Ueno S.-I., Kakimoto Y.,
RA Morita N.;
```

```
RT "Saposin-C from bovine spleen; complete amino acid sequence and
RT relation between the structure and its biological activity.";
RL Biochim. Biophys. Acta 1120:75-80(1992).
CC -i- FUNCTION: The lysosomal degradation of sphingolipids takes place
CC by the sequential action of specific hydrolases. Some of these
CC enzymes require specific low-molecular masses, non-enzymic proteins:
CC the sphingolipids activator proteins (coproteins) (By similarity).
CC -i- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC galactosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
CC Saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate.
CC -i- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
CC cerebroside sulfate by arylsulfatase A (EC 3.1.1.6.8), GMI
CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
CC Saposin B forms a solubilizing complex with the substrates of the
CC sphingolipid hydrolases (By similarity).
CC -i- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
CC activator (EC 3.1.4.12) (By similarity).
CC -i- SUBUNIT: Saposin B is a homodimer (By similarity).
CC -i- SUBCELLULAR LOCATION: Lysosomal.
CC -i- PTM: This precursor is proteolytically processed to 4 small
CC peptides, which are similar to each other and are sphingolipid
CC hydrolase activator proteins (By similarity).
CC -i- SIMILARITY: Contains 2 saposin A-type domains.
CC -i- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AB036791; BAA95677.1; -.
DR InterPro; IPR003119; SapA.
DR InterPro; IPR007856; SapB 1.
DR InterPro; IPR008138; SapB 2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SapB 1; 4.
DR Pfam; PF03489; SapB 2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 4.
KW Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat.
FT SIGNAL 1 16
FT PROPEP 17 58
FT CHAIN 60 142 SAPOSIN A.
FT PROPEP 144 195
FT CHAIN 196 275 SAPOSIN B.
FT PROPEP 277 310
FT CHAIN 312 392 SAPOSIN C.
FT PROPEP 393 404
FT CHAIN 406 487 SAPOSIN D.
FT PROPEP 489 525
FT DOMAIN 21 54
FT DOMAIN 59 142 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 194 276 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 312 393 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 406 487 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 492 525 SAPOSIN-LIKE TYPE B 4.
FT DISULFID 63 138 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 199 272 BY SIMILARITY.
FT DISULFID 202 266 BY SIMILARITY.
```

```

FT DISULFID 231 242 BY SIMILARITY.
FT DISULFID 316 389 BY SIMILARITY.
FT DISULFID 319 383 BY SIMILARITY.
FT DISULFID 347 358 BY SIMILARITY.
FT DISULFID 410 483 BY SIMILARITY.
FT DISULFID 413 477 BY SIMILARITY.
FT DISULFID 441 452 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 427 427 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 127 127 H -> R.
FT VARIANT 260 263 MOPK -> IRIH.
FT CONFLICT 317 317 E -> Q (IN REF. 2).
FT CONFLICT 367 367 R -> S (IN REF. 1).
SQ SEQUENCE 525 AA; 58120 MW; 293AFC0F9C4FA99 CRC64;

Query Match 5.7%; Score 30; DB 1; Length 525;
Best Local Similarity 100.0%; Pred. No. 8.7e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 PEVCNALNCESLOKHLAELNHQKLESN 157
DB 128 PEVCNALNCESLOKHLAELNHQKLESN 157

RESULT 3
SAP_PIG STANDARD; PRT; 80 AA.
AC P81405;

DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Saposin B (Cerebroside sulfate activator) (CS-ACT) (Non-specific
DE activator) (Sphingolipid activator protein 1) (SAP-1).
OS Sus scrofa (Fig.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE OF 1-79.
RC TISSUE=Kidney;
RX MEDLINE=92229506; PubMed=8471613;
RA Stevens R.L., Faull K.F., Conklin K.A., Green B.N., Fluhrarty A.L.;
RT "Porcine cerebroside sulfate activator; further structural
RT characterization and disulfide identification.";
RL Biochemistry 32:4051-4059(1993).
RN [2]
RP SEQUENCE OF 1-64.
RC TISSUE=Kidney;
RX MEDLINE=92222651; PubMed=1562358;
RA Fluhrarty A.L., Katona Z., Meek W.E., Frei K., Fowler A.V.;
RT "The cerebroside sulfate activator from pig kidney: purification and
RT molecular structure.";
RL Biochem. Med. Metab. Biol. 47:66-85(1992).
RN [3]
RP STRUCTURE OF CARBOHYDRATE ON ASN-21.
RX MEDLINE=21110404; PubMed=11180632;
RA Faull K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,
RA Stevens R.L., Fluhrarty C.B., Fluhrarty A.L.;
RT "Structure of the asparagine-linked sugar chains of porcine kidney and
RT human urine cerebroside sulfate activator protein.";
RL J. Mass Spectrom. 35:1416-1424(2000).
RN [4]
RP MASS SPECTROMETRY.
RC TISSUE=Kidney;
RX MEDLINE=99441404; PubMed=10510427;
RA Faull K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,
RA Krutchinsky A.N., Standing K.G., Waring A.J., Stevens R.L.,
RA Fluhrarty C.B., Fluhrarty A.L.;
RT "Cerebroside sulfate activator protein (Saposin B): chromatographic
RT and electrospray mass spectrometric properties.";

```

```

RL J. Mass Spectrom. 34:1040-1054(1999).
CC -1- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GMI
CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
CC Saposin B forms a solubilizing complex with the substrates of the
CC sphingolipid hydrolases.
CC -1- SUBUNIT: Saposin B is a homodimer (By similarity).
CC -1- PTM: The one residue extended Saposin B-Val is only found in a
CC minority of the chains.
CC -1- SIMILARITY: Contains 1 saposin B-type domain.
DR GlycoSuiteDB; P81405; -.
DR InterPro; IPR007856; SapB 1.
DR InterPro; IPR008138; SapB 2.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR01797; SAPOSIN.
DR SMART; SM00118; SAPB; 1.
KW Glycoprotein; Sphingolipid metabolism.
FT CHAIN 1 79
FT CHAIN 1 80 SAPOSIN B.
FT DOMAIN 1 80 SAPOSIN B-VAL.
FT DISULFID 4 77 SAPOSIN-LIKE TYPE B.
FT DISULFID 7 71
FT DISULFID 36 47
FT CARBOHYD 21 21
FT SEQUENCE 80 AA; 8949 MW; EF7BA249B63E789C CRC64;

Query Match 4.2%; Score 22; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.8e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 CKNVISQYSEIAIQMMHMQPK 262
DB 47 CKNVISQYSEIAIQMMHMQPK 69

RESULT 4
SAP_RAT STANDARD; PRT; 554 AA.
AC P10560; Q62841; Q64190;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN PSAP OR SGP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Sertoli cells;
RX MEDLINE=89000647; PubMed=3048385;
RA Collard M.W., Sylvester S.R., Tsuruta J.K., Griswold M.D.;
RT "Biosynthesis and molecular cloning of sulfated glycoprotein 1
RT secreted by rat Sertoli cells; sequence similarity with the
RL 70-kilodalton precursor to sulfatide/GM1 activator.";
RN [2]
RL Biochemistry 27:4557-4564(1988).
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96128541; PubMed=8573994;
RA Morales C.R., El-Alfy M., Zhao Q., Igodura S.A.;
RT "Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in
RL Sertoli cells.";
RN [3]
RN Histol. Histopathol. 10:1023-1034(1995).
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;

```

```
RX MEDLINE=96175245; PubMed=8601692;
RA Morales C.R., El-Alfy M., Zhao Q., Igdoura S.A.;
RT "Expression and tissue distribution of rat sulfated glycoprotein-1
  (prosaposin).";
RL J. Histochem. Cytochem. 44:327-337(1996).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: Contains 2 saposin A-type domains.
CC -!- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19936; AAA42136.1; -
DR EMBL; S81353; AAB36042.2; -
DR EMBL; S81373; AAB36233.2; -
DR PIR; A28716; A28716.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SapB_1; 3.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 4.
DR Signal; Glycoprotein; Repeat.
FT SIGNAL 1 16
FT CHAIN 17 554 SULFATED GLYCOPROTEIN 1.
FT DOMAIN 21 54 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 61 138 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 193 274 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 310 391 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 435 516 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 521 554 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 197 270 BY SIMILARITY.
FT DISULFID 200 264 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 314 387 BY SIMILARITY.
FT DISULFID 317 381 BY SIMILARITY.
FT DISULFID 345 356 BY SIMILARITY.
FT DISULFID 439 512 BY SIMILARITY.
FT DISULFID 442 506 BY SIMILARITY.
FT DISULFID 470 481 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 115 115 P -> L (IN REF. 2 AND 3).
FT CONFLICT 299 299 D -> E (IN REF. 2).
FT CONFLICT 462 462 I -> V (IN REF. 3).
FT CONFLICT 527 527 W -> R (IN REF. 3).
FT CONFLICT 536 536 S -> M (IN REF. 3).
SQ SEQUENCE 554 AA; 61123 MW; DFE3F3A3A0520C6B CRC64;
Query Match 4.2%; Score 22; DB 1; Length 554;
Best Local Similarity 100.0%; Pred. No. 9.6e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 431 EILAALEKGCFLPDPYKQCD 452
Db 461 EILAALEKGCFLPDPYKQCD 482
```

```
RESULT 5
SAP_MOUSE STANDARD; PRT; 557 AA.
ID Q61207; Q60861; Q64006; Q64219;
AC Q61207; Q60861; Q64006; Q64219;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN PSAP OR SGPI
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92272718; PubMed=1590788;
RA Teuda M., Sakiyama T., Endo H., Kitagawa T.;
RT "The primary structure of mouse saposin.";
RL Biochem. Biophys. Res. Commun. 184:1266-1272(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94272317; PubMed=8003952;
RA Sprecher-Levy H., Orr-Urtreger A., Lonai P., Horowitz M.;
RT "Murine prosaposin: expression in the reproductive system of a gene
  implicated in human genetic disease.";
RL Cell. Mol. Biol. 40:233-233(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96084310; PubMed=8565332;
RA Cao Q.P., Crain W.R.;
RT "Expression of SGP-1 mRNA in preimplantation mouse embryos.";
RL Dev. Genet. 17:263-271(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Zhao Q.Q., Hay N.N., Morales C.R.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: Contains 2 saposin A-type domains.
CC -!- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S36200; AAB22175.1; -
DR EMBL; S71616; AAB31059.1; -
DR EMBL; U27340; AAA92567.1; -
DR PIR; JH0604; JH0604.
DR MGI; MGI:97783; Psap.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SapB_1; 4.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 4.
DR Signal; Glycoprotein; Repeat.
FT SIGNAL 1 16 BY SIMILARITY.
```

FT CHAIN 17 557
 FT DOMAIN 21 54
 FT DOMAIN 59 142
 FT DOMAIN 193 277
 FT DOMAIN 313 394
 FT DOMAIN 438 519
 FT DOMAIN 524 557
 FT DISULFID 63 138
 FT DISULFID 66 132
 FT DISULFID 94 106
 FT DISULFID 197 273
 FT DISULFID 200 267
 FT DISULFID 229 240
 FT DISULFID 317 390
 FT DISULFID 320 384
 FT DISULFID 348 359
 FT DISULFID 442 515
 FT DISULFID 445 509
 FT DISULFID 473 484
 FT CARBOHYD 80 80
 FT CARBOHYD 214 214
 FT CARBOHYD 334 334
 FT CARBOHYD 459 459
 FT CONFLICT 83 83
 FT CONFLICT 158 158
 FT CONFLICT 160 160
 FT CONFLICT 171 172
 FT CONFLICT 244 244
 FT CONFLICT 254 254
 FT CONFLICT 255 255
 FT CONFLICT 260 262
 FT CONFLICT 307 307
 FT CONFLICT 322 322
 FT CONFLICT 349 350
 FT CONFLICT 367 367
 FT CONFLICT 370 370
 FT CONFLICT 373 373
 FT CONFLICT 391 391
 FT CONFLICT 393 393
 FT CONFLICT 406 406
 FT CONFLICT 430 430
 FT CONFLICT 445 445
 FT CONFLICT 448 448
 SQ SEQUENCE 557 AA; 61422 MW; 134593E20499535E CRC64;

Query Match 4.2%; Score 22; DB 1; Length 557;
 Best Local Similarity 100.0%; Pred. No. 9.7e-14;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 431 EILAALEKGCFLDPDPYKQCD 452
 Db 464 EILAALEKGCFLDPDPYKQCD 485

RESULT 6
 ID -SAP CHICK STANDARD; PRT; 518 AA.
 AC O13035;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Proactivator polypeptide precursor [Contains: Saposin A; Saposin B;
 DE Saposin C; Saposin D].
 GN PSAP.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 194-203.
 RC TISSUE=Brain, and Liver;
 RX MEDLINE=98129745; PubMed=9461526;

RA Azuma N., Seo H.-C., Ito O., Fu Q., Gould R.M., Hiraiwa M., Burt D.W.,
 RA Paton I.R., Morrice D.R., O'Brien J.S., Kishimoto Y.,
 RT "Cloning, expression and map assignment of chicken prosaposin.";
 RL Biochem. J. 330:321-327(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Altman N., Horowitz M.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: The lysosomal degradation of sphingolipids takes place
 CC by the sequential action of specific hydrolases. Some of these
 CC enzymes require specific low-molecular mass, non-enzymic proteins:
 CC the sphingolipids activator proteins (coproteins) (By similarity).
 CC -!- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
 CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
 CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
 CC Saposin C apparently acts by combining with the enzyme and acidic
 CC lipid to form an activated complex, rather than by solubilizing
 CC the substrate (By similarity).
 CC -!- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
 CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1
 CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
 CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
 CC Saposin B forms a solubilizing complex with the substrates of the
 CC sphingolipid hydrolases (By similarity).
 CC -!- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
 CC activator (EC 3.1.4.12) (By similarity).
 CC -!- SUBUNIT: Saposin B is a homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Lysosomal (By similarity).
 CC -!- PTM: This precursor is proteolytically processed to 4 small
 CC peptides, which are similar to each other and are sphingolipid
 CC hydrolase activator proteins (By similarity).
 CC -!- SIMILARITY: Contains 2 saposin A-type domains.
 CC -!- SIMILARITY: Contains 4 saposin B-type domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB003471; BAA19914.1; -;
 DR EMBL; AF108656; AAF05899.1; -;
 DR InterPro; IPR003119; Sapa.
 DR InterPro; IPR007856; SapaB_1.
 DR InterPro; IPR008138; SapaB_2.
 DR InterPro; IPR008140; SapaB_sub.
 DR InterPro; IPR008373; Saposin.
 DR InterPro; IPR008139; SaposinB.
 DR Pfam; PF02199; SAPA; 2.
 DR Pfam; PF05184; SapaB_1; 4.
 DR Pfam; PF03489; SapaB_2; 4.
 DR PRINTS; PR01797; SAPOSIN.
 DR ProDom; PD001732; SapaB_sub; 3.
 KW Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat;
 KW GM2-gangliosidosis.
 FT SIGNAL 1 17 POTENTIAL.
 FT PROPEP 18 60
 FT CHAIN 61 143 SAPOSIN A.
 FT PROPEP 145 193
 FT CHAIN 194 276 SAPOSIN B.
 FT PROPEP 278 305
 FT CHAIN 307 387 SAPOSIN C.
 FT PROPEP 389 398
 FT CHAIN 399 480 SAPOSIN D.
 FT PROPEP 482 518
 FT DOMAIN 22 55
 FT DOMAIN 60 143 SAPOSIN-LIKE TYPE A 1.
 FT DOMAIN 193 277 SAPOSIN-LIKE TYPE B 1.
 FT DOMAIN 307 388 SAPOSIN-LIKE TYPE B 2.
 FT DOMAIN 399 480 SAPOSIN-LIKE TYPE B 3.
 FT DOMAIN 485 518 SAPOSIN-LIKE TYPE B 4.
 FT DOMAIN 485 518

FT DISULFID 64 139 BY SIMILARITY.
 FT DISULFID 67 133 BY SIMILARITY.
 FT DISULFID 95 107 BY SIMILARITY.
 FT DISULFID 197 273 BY SIMILARITY.
 FT DISULFID 200 267 BY SIMILARITY.
 FT DISULFID 229 240 BY SIMILARITY.
 FT DISULFID 311 384 BY SIMILARITY.
 FT DISULFID 314 378 BY SIMILARITY.
 FT DISULFID 342 353 BY SIMILARITY.
 FT DISULFID 403 476 BY SIMILARITY.
 FT DISULFID 406 470 BY SIMILARITY.
 FT DISULFID 434 445 BY SIMILARITY.
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 94 94 R -> T (IN REF. 2).
 FT CONFLICT 486 486 E -> D (IN REF. 2).
 SQ SEQUENCE 518 AA; 57601 MW; B803000E891C3963 CRC64;

Query Match 2.3%; Score 12; DB 1; Length 518;
 Best Local Similarity 100.0%; Pred. No. 0.00096;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 CKHIVDSYLPVI 117
 Db 107 CKHIVDSYLPVI 118
 |||||

RESULT 7
 SAP_CAVPO STANDARD; PRT; 81 AA.
 AC P20097;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Saposin C (CO-beta-glucosidase) (Glucosylceramidase activator)
 DE (Sphingolipid activator protein 2) (SAP-2).
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=89066787; PubMed=3198642;
 RA Sano A., Radin N.S., Johnson L.L., Tarr G.E.;
 RT "The activator protein for glucosylceramide beta-glucosidase from
 RT guinea pig liver. Improved isolation method and complete amino acid
 RT sequence."
 RL J. Biol. Chem. 263:19597-19601(1988).
 CC -1- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
 CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
 CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
 CC Saposin C apparently acts by combining with the enzyme and acidic
 CC lipid to form an activated complex, rather than by solubilizing
 CC the substrate.
 CC -1- SIMILARITY: Contains 1 saposin B-type domain.
 DR PIR: A32026; A32026.
 DR HSP; P42210; IQDM.
 DR InterPro: IPR007856; SapB_1.
 DR InterPro: IPR008138; SapB_2.
 DR InterPro: IPR008140; SapB_sub.
 DR InterPro: IPR008173; Saposin.
 DR InterPro: IPR008139; SaposinB.
 DR Pfam: PF05184; SapB_1; 1.
 DR Pfam: PF03489; SapB_2; 1.
 DR PRINTS; PRO1797; SAPOSIN.
 DR ProDom; PD001732; SapB_sub; 1.
 DR SMART; SM00118; SAPB; 1.
 KW Glycoprotein; Sphingolipid metabolism.
 FT DOMAIN 1 81 SAPOSIN-LIKE TYPE B.
 FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .).

FT DISULFID 5 78 BY SIMILARITY.
 FT DISULFID 8 72 BY SIMILARITY.
 FT DISULFID 36 47 BY SIMILARITY.
 SQ SEQUENCE 81 AA; 8852 MW; E564CE1FOA292596 CRC64;
 Query Match 1.7%; Score 9; DB 1; Length 81;
 Best Local Similarity 100.0%; Pred. No. 0.19;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 357 CQEVDTYG 365
 Db 47 CQEVDTYG 55
 |||||
 RESULT 8
 FIMI_ECOLI STANDARD; PRT; 179 AA.
 AC P39264;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Fimbrin-like protein fimi precursor.
 GN FIMI OR B4315.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=95334362; PubMed=7610040;
 RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
 RA Blattner F.R.;
 RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
 RT region from 92.8 through 100 minutes."
 RL Nucleic Acids Res. 23:2105-2119(1995).
 RN [2]
 RP PRELIMINARY SEQUENCE OF 1-80 FROM N.A.
 RX MEDLINE=84285425; PubMed=6147250;
 RA Klemm P.;
 RT "The fima gene encoding the type-1 fimbrial subunit of Escherichia
 RT coli. Nucleotide sequence and primary structure of the protein."
 RL Bur. J. Biochem. 143:395-399(1984).
 CC -1- SIMILARITY: BELONGS TO THE FIMA/PAPA FAMILY OF FIMBRIA PROTEINS.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to
 CC frameshifts.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U14003; AAA97211.1; ALT INIT.
 CC EMBL; AE000502; AAC77271.1; ALT INIT.
 CC EMBL; X00981; -: NOT ANNOTATED_CDS.
 CC EcoGene; EGI1974; fimi.
 DR InterPro; IPR008966; Adhes_bact.
 DR InterPro; IPR000259; Fimbrial.
 DR Pfam; PF00419; Fimbrial; 1.
 KW Fimbria; Signal; Complete proteome.
 FT SIGNAL 1 19 POTENTIAL
 FT CHAIN 20 179 FIMBRIN-LIKE PROTEIN FIMI.
 FT DISULFID 43 83 BY SIMILARITY.
 SQ SEQUENCE 179 AA; 19373 MW; ECE4FA7D24B85E25 CRC64;
 Query Match 1.5%; Score 8; DB 1; Length 179;
 Best Local Similarity 100.0%; Pred. No. 3.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 4 LFLASLL 11

```

Db 6 LFLASL 13
|||||
RESULT 9
HS23 DROME STANDARD; PRT; 186 AA.
AC P02516; Q9V5X5;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Heat shock protein 23.
GN HSP23 OR CG4463.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83189140; PubMed=6302284;
RA Southgate R., Ayme A., Voellmy R.;
RT "Nucleotide sequence analysis of the Drosophila small heat shock gene
cluster at locus 67B.";
RL J. Mol. Biol. 165:35-57(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=82248004; PubMed=6285380;
RA Ingolia T.D., Craig E.A.;
RT "Four small Drosophila heat shock proteins are related to each other
and to mammalian alpha-crystallin.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2360-2364(1982).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Bueam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleeb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";

```

```

RL Science 287:2185-2195(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Krommiller B., Pacleeb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -I- SIMILARITY: Belongs to the small heat shock protein (HSP20)
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J01100; AAA28637.1; -
CC EMBL; V00210; CAA23494.1; -
CC EMBL; X03889; CAA27525.1; -
CC EMBL; A2003552; AAF50286.1; -
CC EMBL; AY061081; AAL28629.1; -
CC PIR; B20647; E20647.
CC Flybase; FBgn0001224; Hsp23.
CC InterPro; IPR001436; Crystallin_alpha.
CC InterPro; IPR002068; Hsp20.
CC InterPro; IPR008978; HSP20_chap.
CC Pfam; PF00011; HSP20; 1.
CC PRINTS; PR002299; ACRYSTALLIN.
CC PROSITE; PS01031; HSP20; 1.
KW Heat shock; Multigene family.
FT CONFLICT 31 33 QRN -> RRI (IN REF. 2).
FT CONFLICT 81 81 K -> E (IN REF. 2).
FT CONFLICT 88 88 K -> G (IN REF. 2).
FT CONFLICT 95 95 L -> V (IN REF. 2).
FT CONFLICT 173 173 N -> S (IN REF. 2).
FT CONFLICT 182 182 N -> G (IN REF. 2).
SQ SEQUENCE 186 AA; 20629 MW; 3C6EDCDAL719BSA0 CRC64;
Query Match 1.5%; Score 8; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 172 MANIPLLL 179
DB 1 MANIPLLL 8
RESULT 10
FMFX BORPE STANDARD; PRT; 201 AA.
AC P09808;
DT 01-MAR-1989 (Rel. 10, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Fimbrial protein fimX precursor (Pilin).
GN FIMX OR BF2674.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sal;
RX MEDLINE=89013896; PubMed=2902506;
RA Pedroni P., Riboli B., de Ferra F., Grandi G., Toma S., Arico B.,
RA Rappuoli R.;
RT "Cloning of a novel pilin-like gene from Bordetella pertussis:
homology to the fim2 gene.";

```



```

RL Mol. Microbiol. 2:539-543 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cordero-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RL Nat. Genet. 35:32-40 (2003).
CC -!- SUBCELLULAR LOCATION: Pili structure on the cell surface.
CC -!- DISEASE: Bordetella pertussis is the causative agent of whooping
CC cough. An essential step in the disease process is the attachment
CC of the bacteria to the ciliated epithelium of the respiratory
CC tract, enabling the organism to resist normal host-clearance
CC mechanisms. It is unclear which bacterial cell surface component
CC are responsible for adherence but the fimbriae of B.pertussis are
CC prime candidates for being involved in this process.
CC -!- SIMILARITY: BELONGS TO THE FIMB/PAPA FAMILY OF FIMBRIA PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y00556; CA68634.1; -.
DR EMBL; AX5464; CAA01770.1; -.
DR EMBL; BX640419; CAB42950.1; -.
DR PIR; S01929; S01929.
DR InterPro; IPR008966; Adhes bact.
DR PIRam; PF00419; Fimbrial; 1.
DR InterPro; IPR00259; Fimbrial.
KW Fimbrin; Signal; Complete proteome.
FT SIGNAL 1 21
FT CHAIN 22 201 FIMBRIAL PROTEIN FIMX.
FT DISULFID 37 79 PROBABLE.
FT CONFLICT 165 166 AK -> GQ (IN REF. 1).
SQ SEQUENCE 201 AA; 21462 MW; F47670CC055B03F6 CRC64;

Query Match 1.5%; Score 8; DB 1; Length 201;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 LLGAALAG 17
Db 7 LLGAALAG 14

RESULT 11
Y726 METJA
AC Q58136;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein WJ0726.
GN MJ0726.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_taxid=2190;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Blutt G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kierlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RN jannaschii."
RL Science 273:1058-1073 (1996).
CC -!- SIMILARITY: SOME, TO COENZYME F420 HYDROGENASE GAMMA SUBUNIT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67519; AAB98722.1; -.
DR PIR; F64390; F64390.
DR TIGR; MJ0726; -.
DR InterPro; IPR001450; 4Fe4S ferredoxin.
DR InterPro; IPR006137; Oxidored_q6.
DR PIRam; PF00037; fer4; 1.
DR PIRam; PF01058; Oxidored_q6; 1.
DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
KW Hypothetical protein; Oxidoreductase; Iron-sulfur; 4Fe-4S;
KW Electron transport; Complete proteome.
FT METAL 169 169 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 172 172 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 175 175 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 179 179 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 197 197 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 200 200 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 203 203 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 207 207 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
SQ SEQUENCE 216 AA; 23789 MW; 2FD9F067698982FE CRC64;

Query Match 1.5%; Score 8; DB 1; Length 216;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 KEIVDSYL 114
Db 117 KEIVDSYL 124

RESULT 12
BOR4 MOUSE
ID BOR4 MOUSE STANDARD; PRT; 349 AA.
AC Q9J36; Q9QZT8;
DT 28-FEB-2003 (Rel. 41, Created).
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cdc42 effector protein 4 (Binder of Rho GTPase 4).
GN CDC42EP4 OR BOR4 OR CEP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=ileal mucosa;
RX MEDLINE=21036164; PubMed=11185749;
RA Osada N., Kusuda J., Suzuki Y., Sugano S., Hashimoto K.;
RT "Sequence analysis, gene expression, and chromosomal assignment of
RN mouse Borg4 gene and its human orthologue."

```

J. Hum. Genet. 45:374-377(2000).
 [2]
 RL SEQUENCE FROM N.A.
 RN MEDLINE=22388257; PubMed=12477932;
 RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan K.C., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., McEwan P.J., McKernan K.J., Abramson R.D., Mullahy S.J.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE OF 1-143 FROM N.A., AND INTERACTION WITH TC10 AND CDC42.
 RX MEDLINE=99421943; PubMed=10490598;
 RA Joberty G., Perlungher R.R., Macata I.G.;
 RT "The Borgs, a new family of Cdc42 and TC10 GTPase-interacting
 proteins";
 RL Mol. Cell. Biol. 19:6585-6597(1999).
 CC -!- FUNCTION: Probably involved in the organization of the actin
 CC cytoskeleton. May act downstream of CDC42 to induce actin
 CC filament assembly leading to cell shape changes. Induces
 CC pseudopodia formation, when overexpressed in fibroblasts.
 CC -!- SUBUNIT: Interacts with CDC42 and TC10 in a GTP-dependent manner.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; membrane-associated (By
 CC similarity).
 CC -!- TISSUE SPECIFICITY: Ubiquitous.
 CC -!- SIMILARITY: BELONGS TO THE BORG/CEP FAMILY.
 CC -!- SIMILARITY: Contains 1 CRIB domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB035088; BAA95932.1; -
 DR EMBL; BC003857; AAH03857.1; -
 DR EMBL; AF165114; AAD47822.1; -
 DR MGD; MGI:1923760; Cdc42ep4.
 DR GO; GO:0005737; Cytoplasm; IDA.
 DR GO; GO:0005515; F-protein binding; IPI.
 DR GO; GO:0007266; P-Rho protein signal transduction; IDA.
 DR Pfam; PF00786; PBD; 1
 DR PROSITE; PSS0108; CRIB; 1.
 DR CYTOSKELETON.
 DR DOMAIN 27 41 CRIB.
 DR CONFLICT 143 143
 DR SEQUENCE 349 AA; 37869 MW; 7C44125A7083E16B CRC64;
 Query Match 1.5%; Score 8; DB 1; Length 349;
 Best Local Similarity 100.0%; Pred. No. 7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 347 SKLPSKLS 354
 DB 131 SKLPSKLS 138

RESULT 13
 BOR4_HUMAN STANDARD; PRT; 356 AA.
 AC Q9H3QJ; Q95828; Q96FT3;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cdc42 effector protein 4 (Binder of Rho GTPase 4).
 GN CDC42EP4 OR BOR4 OR CEP4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21036164; PubMed=11185749;
 RA Osada N., Kusuda J., Suzuki Y., Sugano S., Hashimoto K.;
 RT "Sequence analysis, gene expression, and chromosomal assignment of
 mouse Borg4 gene and its human orthologue";
 RL J. Hum. Genet. 45:374-377(2000).
 RN [2]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX TISSUE=Eye;
 MEDLINE=22125609; PubMed=11035016;
 RA Hirsch D.S., Pirone D.M., Burbelo P.D.;
 RT "A new family of cdc42 effector proteins, CEPs, function in fibroblast
 and epithelial cell shape changes";
 RL J. Biol. Chem. 276:875-883(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Brain;
 MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan K.C., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., McEwan P.J., McKernan K.J., Abramson R.D., Mullahy S.J.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Probably involved in the organization of the actin
 CC cytoskeleton. May act downstream of CDC42 to induce actin
 CC filament assembly leading to cell shape changes. Induces
 CC pseudopodia formation, when overexpressed in fibroblasts.
 CC -!- SUBUNIT: Interacts with CDC42 and TC10 in a GTP-dependent manner
 CC (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; membrane-associated.
 CC -!- TISSUE SPECIFICITY: Not detected in any of the adult tissues
 CC tested. May be expressed only in fetal or embryonic tissues.
 CC -!- SIMILARITY: BELONGS TO THE BORG/CEP FAMILY.
 CC -!- SIMILARITY: Contains 1 CRIB domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```
CC -----
DR EMBL; AB042237; BAB1272.1; -
DR EMBL; AF09664; AAD16299.1; -
DR EMBL; BC010451; AAH10451.1; -
DR Genbank; HGNC:17147; CDC42EP4.
DR MM; 605468; -
DR InterPro; IPR000095; PAKbox/RhoBndng.
DR Pfam; PF00786; PBD; 1.
DR PROSITE; PS0108; CRIB; 1.
KW Cytoskeleton.
FT DOMAIN 27 41 CRIB.
FT DOMAIN 70 74 POLY-SER.
FT CONFLICT 2 2 P -> L (IN REF. 3).
FT CONFLICT 33 33 D -> T (IN REF. 2).
FT CONFLICT 288 288 A -> T (IN REF. 2).
FT CONFLICT 296 296 A -> T (IN REF. 2).
FT CONFLICT 339 339 P -> PP (IN REF. 2).
FT CONFLICT 355 355 MISSING (IN REF. 2).
SQ SEQUENCE 356 AA; 37979 MW; 2CF677C60C6EF1B5 CRC64;

Query Match 1.5%; Score 8; DB 1; Length 356;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 347 SKLPKSL 354
Db 133 SKLPKSL 140
|||||

RESULT 14
TCR3_ECOLI ID TCR3_ECOLI STANDARD; PRT; 396 AA.
AC P02981;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tetracycline resistance protein, class C (TETA(C)).
GN TETA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80002802; PubMed=383387;
RA Sutcliffe J.G.;
RT "Complete nucleotide sequence of the Escherichia coli plasmid pBR322.";
RL Cold Spring Harb. Symp. Quant. Biol. 43:77-90(1979).
RN [2]
RP REVISIONS, AND IDENTIFICATION OF PROTEIN.
RX MEDLINE=83117828; PubMed=6337373;
RA Livneh Z.;
RT "Directed mutagenesis method for analysis of mutagen specificity: application to ultraviolet-induced mutagenesis.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:237-241(1983).
RN [3]
RP REVISIONS, AND IDENTIFICATION OF PROTEIN.
RX MEDLINE=83263146; PubMed=6307828;
RA Peden K.W.C.;
RT "Revised sequence of the tetracycline-resistance gene of pBR322.";
RL Gene 22:277-280(1983).
CC -1- FUNCTION: Resistance to tetracycline by an active tetracycline efflux. This is an energy-dependent process that decreases the accumulation of the antibiotic in whole cells. This protein functions as a metal-tetracycline/H(+) antiporter.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSLUCASE FAMILY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation
```

```
CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J01749; AAB59735.1; -
DR PIR; B90923; YTEC32.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR001958; TCR_Teta.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR01035; TCR_TETA.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN 1.
KW Antibiotic resistance; Transmembrane; Inner membrane; Transport;
KW Antiport; Ion transport; Hydrogen ion transport; Plasmid.
FT TRANSMEM 8 28 POTENTIAL.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 104 124 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 161 181 POTENTIAL.
FT TRANSMEM 216 236 POTENTIAL.
FT TRANSMEM 246 267 POTENTIAL.
FT TRANSMEM 279 299 POTENTIAL.
FT TRANSMEM 300 320 POTENTIAL.
FT TRANSMEM 338 358 POTENTIAL.
FT TRANSMEM 365 385 POTENTIAL.
SQ SEQUENCE 396 AA; 41510 MW; 0B9C82B811A2F7DF CRC64;

Query Match 1.5%; Score 8; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LLASLLGA 13
Db 76 LLASLLGA 83
|||||

RESULT 15
MNTC_BACSU ID MNTC_BACSU STANDARD; PRT; 435 AA.
AC O35024;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Manganese transport system membrane protein mntC.
GN MNTC OR BSU30750.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes in the 200 kb rmb-dnaB region.";
RL Microbiology 143:3431-3441(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.W., Dusterhoft A., Ehrlich S.D., Emerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
```

Search completed: May 5, 2004, 13:46:14
Job time : 11.5605 secs

RA Ghim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hoono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Kjaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Toso V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*,"
RN Nature 390:249-256(1997).
RN [3]
RP POSSIBLE FUNCTION.
RX MEDLINE=20223631; PubMed=10760146;
RA Que O., Helmann J.D.;
RT "Manganese homeostasis in *Bacillus subtilis* is regulated by MntR, a
RT bifunctional regulator related to the diphtheria toxin repressor
RL Mol Microbiol. 35:1454-1468(2000).
CC -!- FUNCTION: This protein is probably a component of a manganese
CC permease, a binding protein-dependent, ATP-driven transport
CC system.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO THE ABC-3 SUBFAMILY OF INTEGRAL MEMBRANE
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF008220; AAC00231.1; -;
DR EMBL; Z99119; CAB15053.1; -;
DR PIR; D69992; D69992.
DR Subtilist; BG13853; mntC.
DR InterPro; IPR001626; ABC_transpt3.
DR Pfam; PF00950; ABC-3; 1.
KW Transport; Ion transport; Transmembrane; Complete proteome.
FT TRANSMEM 17 37 POTENTIAL.
FT TRANSMEM 42 62 POTENTIAL.
FT TRANSMEM 68 88 POTENTIAL.
FT TRANSMEM 98 118 POTENTIAL.
FT TRANSMEM 143 163 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT TRANSMEM 228 248 POTENTIAL.
FT TRANSMEM 255 275 POTENTIAL.
SQ SEQUENCE 435 AA; 47944 MW; 68E6D590D90A73BC CRC64;

Query Match 1.5%; Score 8; DB 1; Length 435;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 LLGAALAG 17
| | | | | | | |
Db 71 LLGAALAG 78

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:42:09 ; Search time 16.595 Seconds
(without alignments)
3037.320 Million cell updates/sec

Title: US-09-743-684A-1
Perfect score: 524
Sequence: 1 MYALFLASLLGAALAGPVL.....NTETAAQCNAVECHKRHVWN 524

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	284	54.2	527	1 SAHUP	saposin precursor
2	22	4.2	79	2 A49475	cerebroside sulfat
3	22	4.2	554	1 A28716	saposin precursor
4	22	4.2	557	1 JH0604	saposin precursor
5	13	2.5	80	2 S21770	saposin-C - bovine
6	9	1.7	81	2 A32026	glucosylceramide b
7	8	1.5	162	2 AD1758	hypothetical prote
8	8	1.5	177	2 T47133	hypothetical prote
9	8	1.5	186	1 HHFF23	heat shock protein
10	8	1.5	186	2 B20647	heat shock protein
11	8	1.5	201	2 S36451	fimbrial protein f
12	8	1.5	201	2 S01929	fimbrial protein p
13	8	1.5	215	2 S56540	fiml protein - Esc
14	8	1.5	215	2 E86129	fimbrial protein (
15	8	1.5	215	2 B91288	fimbrial protein (
16	8	1.5	216	2 T64390	coenzyme F420 hydr
17	8	1.5	216	2 T48482	hypothetical prote
18	8	1.5	225	2 T02241	probable germin pr
19	8	1.5	274	4 JC5585	hypothetical mutan
20	8	1.5	326	2 B90841	probable ABC trans
21	8	1.5	326	2 A85699	probable ABC trans
22	8	1.5	354	2 AF3614	probable undecapre
23	8	1.5	396	1 YTEC32	tetracycline resis
24	8	1.5	435	2 D69992	ABC transporter (m
25	8	1.5	655	2 D85988	hypothetical prote
26	8	1.5	655	2 B65116	hypothetical prote
27	8	1.5	655	2 A91143	hypothetical prote
28	8	1.5	813	1 A49123	fibroblast growth
29	7	1.3	62	2 S77874	probable preprotei

RESULT 1 SAHUP

saposin precursor [validated] - human

N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component
ein (SAP); sphingolipid activator protein A2; sulfate sulfatase activator protein
N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1992 #sequence_revision 17-Nov-1995 #text_change 08-Dec-2000
C;Accession: JX0061; A57368; A42003; B42003; C42003; D42003; A30367; S34740; S36140; S36
0226; I37265; I37264
J;Nakano, T.; Sandhoff, K.; Stuenkel, J.; Christomanou, H.; Suzuki, K.
J;Biochem. 105, 152-154, 1989
A;Title: Structure of full-length cDNA coding for sulfate activator, a Co-beta-glucosidase
A;Reference number: JX0061; MUID:89255151; PMID:2498298
A;Accession: JX0061
A;Molecule type: mRNA
A;Residues: 1-527 <NAK>
A;Cross-references: GB:D00422; NID:g220063; PIDN:BAA00321.1; PID:g220064
A;Note: alternative splice form 1
A;Accession: A57368
A;Molecule type: mRNA
A;Residues: 1-259,263-527 <NA2>
A;Cross-references: GB:J03015; GB:J03086; NID:g337755; PIDN:AAB59494.1; PID:g337756
A;Note: alternative splice form 2
R;Rorman, E.G.; Scheinker, V.; Grabowski, G.A.
Genomics 13, 312-318, 1992
A;Title: Structure and evolution of the human prosaposin chromosomal gene.
A;Reference number: A42003; MUID:92307663; PMID:1612590
A;Accession: A42003
A;Molecule type: DNA
A;Residues: 50-140 <ROR>
A;Cross-references: GB:M86181
A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107236)
A;Accession: B42003
A;Molecule type: DNA
A;Residues: 185-259;263-276 <RO2>
A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107237)
A;Accession: C42003
A;Molecule type: DNA
A;Residues: 305-393 <RO3>
A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107238); sequence inc
A;Accession: D42003
A;Molecule type: DNA
A;Residues: 399-487 <RO4>
A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107239); sequence inc
R;Rorman, E.G.; Grabowski, G.A.
Genomics 5, 486-492, 1989
A;Title: Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphin
A;Reference number: A30367; MUID:90129043; PMID:2515150
A;Accession: A30367
A;Molecule type: mRNA
A;Residues: 1-259,263-527 <RO5>

hypothetical prote
ribosomal protein
lens fiber membran
chorismate mutase
hypothetical prote
helix-loop-helix p
hypothetical prote
acetylactate synth
transcription regu
probable periplasm
hypothetical prote
helix-loop-helix p
helix-loop-helix p
NADH2 dehydrogenas
DNA-binding protei
uncharacterized me

A;Cross-references: GB:J03077; NID:gl83230; PIDN:AAA52560.1; PID:gl83231
 A;Note: alternative splice form 2
 R;Hiraiwa, M.; O'Brien, J.S.; Kishimoto, Y.; Galdzicka, M.; Fluharty, A.L.; Ginns, E.I.; Arch. Biochem. Biophys. 304, 110-116, 1993
 A;Title: Isolation, characterization, and proteolysis of human prosaposin, the precursor
 A;Reference number: S34740; MUID:93311991; PMID:8323276
 A;Accession: S34740
 A;Molecule type: protein
 A;Residues: 17-24;165-172;180-189;301-305 <HIR>
 R;Tyynelae, J.; Palmer, D.N.; Baumann, M.; Haltia, M. FEBS Lett. 330, 8-12, 1993
 A;Title: Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis.
 A;Reference number: S36140; MUID:93380576; PMID:8370464
 A;Accession: S36140
 A;Molecule type: protein
 A;Residues: 'XX', 62, 'X', 64-65, 'X', 67-79, 'X', 81-84 <TY>
 A;Note: saposin A
 A;Accession: S36141
 A;Molecule type: protein
 A;Residues: 'XX', 413-414, 'X', 416-428, 'X', 430-434 <TV2>
 A;Note: saposin D
 R;Holtzschmidt, H.; Sandhoff, K.; Kwon, H.Y.; Harzer, K.; Nakano, T.; Suzuki, K. J. Biol. Chem. 266, 7556-7560, 1991
 A;Title: Sulfatide activator protein. Alternative splicing that generates three mRNAs and
 A;Reference number: S36988; MUID:91210267; PMID:2019586
 A;Accession: S36988
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-240, 'S', 242-527 <HOL>
 A;Cross-references: EMBL:M60255; NID:g337759; PIDN:AAA36594.1; PID:g337760
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
 A;Note: cerebroside sulfate activator protein mutant MU-9; corresponds to alternative splicing
 A;Accession: S36989
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-240, 'S', 242-259, 263-527 <HO2>
 A;Cross-references: EMBL:M60257; NID:g337764; PIDN:AAA36595.1; PID:g337765
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
 A;Note: cerebroside sulfate activator protein mutant MU-0; corresponds to alternative splicing
 A;Accession: S36990
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-240, 'S', 242-259, 261-527 <HO3>
 A;Cross-references: EMBL:M60258; NID:g337766; PIDN:AAA36596.1; PID:g337767
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
 A;Note: cerebroside sulfate activator protein mutant MU-6; corresponds to alternative splicing
 R;Kondoh, K.; Hinenio, T.; Sano, A.; Kakimoto, Y. Biochem. Biophys. Res. Commun. 181, 286-292, 1991
 A;Title: Isolation and characterization of prosaposin from human milk.
 A;Reference number: PS0330; MUID:92068206; PMID:1958198
 A;Accession: PS0330
 A;Molecule type: protein
 A;Residues: 17-24, 'X', 26 <KON>
 A;Experimental source: milk
 R;Kretz, K.A.; Carson, G.S.; Morimoto, S.; Kishimoto, Y.; Fluharty, A.L.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 87, 2541-2544, 1990
 A;Title: Characterization of a mutation in a family with saposin B deficiency: a glycosylation
 A;Reference number: A35985; MUID:90207231; PMID:2320574
 A;Accession: A35985
 A;Molecule type: mRNA
 A;Residues: 213-221 <KRE>
 A;Cross-references: GB:M32221
 A;Accession: B35985
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-259, 263-527 <KR2>
 A;Cross-references: GB:M32221; NID:g337761; PIDN:AAA60303.1; PID:g337762
 A;Experimental source: lymphoblast
 A;Accession: C35985
 A;Molecule type: mRNA
 A;Residues: 213-216, 'I', 218-221 <KR3>
 A;Note: sequence from patients with activator-deficient metachromatic leukodystrophy; th
 R;Fuerst, W.; Schubert, J.; Machleidt, W.; Meyer, H.B.; Sandhoff, K.

Eur. J. Biochem. 192, 709-714, 1990
 A;Title: The complete amino-acid sequences of human ganglioside GM2 activator protein and
 A;Reference number: S13195; MUID:91006185; PMID:2209618
 A;Accession: S13196
 A;Molecule type: protein
 A;Residues: 195-259, 263-277 <FUE>
 R;Morimoto, S.; Martin, B.M.; Yamamoto, Y.; Kretz, K.A.; O'Brien, J.S.; Kishimoto, Y. Proc. Natl. Acad. Sci. U.S.A. 86, 3389-3393, 1989
 A;Title: Saposin A: second cerebroside activator protein.
 A;Reference number: A32784; MUID:89240739; PMID:2717620
 A;Accession: A32784
 A;Molecule type: protein
 A;Residues: 60-84;86-107;109-119;125-134 <MOR>
 R;O'Brien, J.S.; Kretz, K.A.; Dewji, N.; Wenger, D.A.; Esch, F.; Fluharty, A.L. Science 241, 1098-1101, 1988
 A;Title: Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic
 A;Reference number: A41240; MUID:88321660; PMID:2842863
 A;Accession: A41240
 A;Molecule type: mRNA
 A;Residues: 'GSSR', 18-259, 263-299, 'D', 301-302, 'D', 304-527 <OAB>
 A;Cross-references: GB:J03086
 R;Dewji, N.N.; Wenger, D.A.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 84, 8652-8656, 1987
 A;Title: Nucleotide sequence of cloned cDNA for human sphingolipid activator protein 1 p
 A;Reference number: S02289; MUID:88068647; PMID:2825202
 A;Accession: S02289
 A;Status: significant sequence differences
 A;Molecule type: mRNA
 A;Cross-references: EMBL:J03015
 A;Note: this sequence corrected by A41240
 R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 369, 1361-1365, 1988
 A;Title: Complete amino-acid sequence of the naturally occurring A(2) activator protein f
 A;Reference number: S02028; MUID:89207118; PMID:3242555
 A;Accession: S02028
 A;Molecule type: protein
 A;Residues: 195-259, 263-276 <KLE>
 R;Fuerst, W.; Machleidt, W.; Sandhoff, K. Biol. Chem. Hoppe-Seyler 369, 317-328, 1988
 A;Title: The precursor of sulfatide activator protein is processed to three different pr
 A;Reference number: S00813; MUID:89000190; PMID:3048308
 A;Accession: S00813
 A;Molecule type: protein
 A;Residues: 410-487 <FU2>
 R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 368, 1571-1578, 1987
 A;Title: Complete amino-acid sequence and carbohydrate content of the naturally occurring
 A;Reference number: S00226; MUID:88163077; PMID:3442600
 A;Accession: S00226
 A;Molecule type: protein
 A;Residues: 314-393 <KL2>
 R;Vaccaro, A.M.; Salvioli, R.; Barca, A.; Tatti, M.; Ciaffoni, F.; Maras, B.; Siciliano, J. Biol. Chem. 270, 9953-9960, 1995
 A;Title: Structural analysis of saposin C and B. Complete localization of disulfide bridge
 A;Reference number: A57297; MUID:95247790; PMID:7730378
 A;Contents: annotation; disulfide bonds; glycosylation
 R;Holtzschmidt, H.; Sandhoff, K.; Fuerst, W.; Kwon, H.Y.; Schnabel, D.; Suzuki, K. FEBS Lett. 280, 267-270, 1991
 A;Title: The organization of the gene for the human cerebroside sulfate activator protei
 A;Reference number: I37264; MUID:91192146; PMID:2013321
 A;Accession: I37265
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 59-125 <RES>
 A;Cross-references: EMBL:X57107; NID:g30234; PIDN:CAA40391.1; PID:g30235
 A;Accession: I37264
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 307-516 <RE2>
 A;Cross-references: EMBL:X57108; NID:g30232; PIDN:CAA40392.1; PID:gl565257
 A;Note: sequence revised relative to PID:g30233 (corrected coding region)
 C;Genetics:

A;Gene: GDB:PSAP; GLBA
A;Cross-references: GDB:120366; OMIM:176801
A;Map position: 10q22.1-10q22.1
A;Introns: 83/3; 338/3; 401/1; 453/3; 480/3
A;Note: defects in this gene may cause variant Gaucher disease, variant Tay-Sachs disease
A;Note: list of introns is incomplete

Query Match 54.2%; Score 284; DB 1; Length 527;
Best Local Similarity 99.4%; Pred. No. 1.7e-284;
Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 MYALFLASLIGALAGVILGKCTRGSAVWCQNVKTASDCGAVKHCLOTVWVKPTVKS 60
Db 1 MYALFLASLIGALAGVILGKCTRGSAVWCQNVKTASDCGAVKHCLOTVWVKPTVKS 60

Qy 61 LPCDICDVVTAAGDMLKDNATEEILVLEKTCMDLKPKNNSASCKEIVDSYLPVILDI 120
Db 61 LPCDICDVVTAAGDMLKDNATEEILVLEKTCMDLKPKNNSASCKEIVDSYLPVILDI 120

Qy 121 IKGMSRPGVCSALNLCESLQKHLAEINHQLESNKIPELDMTEVVAPFMANIPLLLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAEINHQLESNKIPELDMTEVVAPFMANIPLLLY 180

Qy 181 PODGPRKPPQKNGDVQCDCIQMTDITQAVRTNSTFVQALVHVKEECDRLGPMADI 240
Db 181 PODGPRKPPQKNGDVQCDCIQMTDITQAVRTNSTFVQALVHVKEECDRLGPMADI 240

Qy 241 CKNYISQYSEITAIOMMHM---QPEICALGVFCDEVKEMPQTLVPAKASKNVIIPALE 297
Db 241 CKNYISQYSEITAIOMMHMQQPEICALGVFCDEVKEMPQTLVPAKASKNVIIPALE 300

Qy 298 LVEPIKKEHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKETILDADFDMCKSLPKSLSEEC 357
Db 301 LVEPIKKEHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKETILDADFDMCKSLPKSLSEEC 360

Qy 358 QEVVDTTGSSTLSILLEVSPVCSLHLCGSTRLPALTVHVTQPKDGGFCFVCKKLVG 417
Db 361 QEVVDTTGSSTLSILLEVSPVCSLHLCGSTRLPALTVHVTQPKDGGFCFVCKKLVG 420

Qy 418 YLDRLNLEKNSKQELIAALEKGCFLPDPYQKQCDQFVAEYEPVLIELVEMDPSFVCL 477
Db 421 YLDRLNLEKNSKQELIAALEKGCFLPDPYQKQCDQFVAEYEPVLIELVEMDPSFVCL 480

Qy 478 KIGACPSAHKPLLTGKTEKIMGPSYWCQNTETAACNAVEHCKRHVN 524
Db 481 KIGACPSAHKPLLTGKTEKIMGPSYWCQNTETAACNAVEHCKRHVN 527

RESULT 2
A49475
cerebroside sulfate activator - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 17-May-1996
C;Accession: A49475
R;Stevens, R.L.; Faull, K.F.; Conklin, K.A.; Green, B.N.; Fluharty, A.L.
Biochemistry 32, 4051-4059, 1993
A;Title: Porcine cerebroside sulfate activator: further structural characterization and
A;Reference number: A49475; MUID:932229506; PMID:8471613
A;Accession: A49475
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-79 <STE>
A;Experimental source: kidney
C;Superfamily: saposin; saposin repeat homology
F;1-79/Domain: saposin repeat homology <SAP>

Query Match 4.2%; Score 22; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 9.1e-15;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 CKNYISQYSEITAIOMMHMQPK 262
|||||

Db 47 CKNYISQYSEITAIOMMHMQPK 68

RESULT 3
A28716
saposin precursor - rat
N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfata
N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A28716
R;Collard, M.W.; Sylvester, S.R.; Tsuruta, J.K.; Griswold, M.D.
Biochemistry 27, 4557-4564, 1988
A;Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat
A;Reference number: A28716; MUID:8900647; PMID:3048385
A;Accession: A28716
A;Molecule type: mRNA
A;Residues: 1-554 <COL>
A;Cross-references: GB:M19936; NID:g206904; PIDN:AAA42136.1; PID:g206905
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
C;Function:
A;Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc
A;Pathway: sphingolipid catabolism
A;Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy
A;Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsul
A;Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiester
C;Superfamily: saposin; saposin repeat homology
C;Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-554/Product: prosaposin #status predicted <PRO>
F;55-148/Domain: saposin repeat homology <SAP1>
F;60-143/Product: saposin A #status predicted <SAP1>
F;189-280/Domain: saposin repeat homology <SAP2>
F;194-273/Product: saposin B #status predicted <SAP1>
F;306-397/Domain: saposin repeat homology <SAP3>
F;310-389/Product: saposin C #status predicted <SAPC>
F;431-522/Domain: saposin repeat homology <SAP4>
F;437-514/Product: saposin D #status predicted <SAPD>
F;63-138,66-132,94-106,439-512,442-506,470-481/Disulfide bonds: #status predicted
F;80,224,331,456/Binding site: carbonylrate (Asn) (covalent) #status predicted
F;197-270,300-264,229-240,314-387,317-381,345-356/Disulfide bonds: #status predicted

Query Match 4.2%; Score 22; DB 1; Length 554;
Best Local Similarity 100.0%; Pred. No. 5.5e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 431 EILAALEKGCFLPDPYQKQCD 452
|||||
Db 461 EILAALEKGCFLPDPYQKQCD 482

RESULT 4
JH0604
saposin precursor - mouse
N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfata
N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: JH0604
R;Tsuda, M.; Sakiyama, T.; Endo, H.; Kitagawa, T.
Biochem. Biophys. Res. Commun. 184, 1266-1272, 1992
A;Title: The primary structure of mouse saposin.
A;Reference number: JH0604; MUID:92272718; PMID:1590788
A;Accession: JH0604
A;Molecule type: mRNA
A;Residues: 1-557 <TSU>
A;Cross-references: GB:S36200; NID:g249386; PIDN:AAB22175.1; PID:g249387
A;Experimental source: liver
C;Function:
A;Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc
A;Pathway: sphingolipid catabolism

A;Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosyl
A;Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsulfa
C;Superfamily: saposin; saposin repeat homology
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-557/Product: prosaposin #status predicted <PRO>
F;55-148/Domain: saposin repeat homology <SAP1>
F;60-143/Product: saposin A #status predicted <SAB1>
F;189-283/Domain: saposin repeat homology <SAP2>
F;194-276/Product: saposin B #status predicted <SAB2>
F;309-400/Domain: saposin repeat homology <SAP3>
F;313-392/Product: saposin C #status predicted <SAPC>
F;434-525/Domain: saposin repeat homology <SAP4>
F;440-517/Product: saposin D #status predicted <SAPD>
F;63-138,66-132,94-106,197-273,200-267,229-240,317-390,320-384,348-359,442-515,445-509,4
F;80,214,334,379,459/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.2%; Score 22; DB 1; Length 557;
Best Local Similarity 100.0%; Pred. No. 5.6e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 EILALEKGCSEFLPDYKQCD 452
|||||
Db 464 EILALEKGCSEFLPDYKQCD 485

RESULT 5
S21770
C;Species: Bos primigenius taurus (cattle)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C;Accession: S21770
R;Sano, A.; Mizuno, T.; Kondoh, K.; Hineno, T.; Ueno, S.; Kakimoto, Y.; Morita, N.
Biochim. Biophys. Acta 1120, 75-80, 1992
A;Title: Saposin-C from bovine spleen: complete amino acid sequence and relation between
A;Reference number: S21770; MUID:92207994; PMID:1554743
A;Accession: S21770
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-80 <SAN>
C;Superfamily: saposin; saposin repeat homology

Query Match 2.5%; Score 13; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 377 SPELVCSMLHCS 389
|||||
Db 67 SPELVCSMLHCS 79

RESULT 6
A32026
Glucosylceramide beta-glucosidase activator protein SAP-2 - guinea pig
C;Species: Cavia porcellus (guinea pig)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 17-May-1996
C;Accession: A32026
R;Sano, A.; Radin, N.S.; Johnson, L.L.; Tarr, G.E.
J. Biol. Chem. 263, 19597-19601, 1988
A;Title: The activator protein for glucosylceramide beta-glucosidase from guinea pig liv
A;Reference number: A32026; MUID:89066787; PMID:3198642
A;Accession: A32026
A;Molecule type: protein
A;Residues: 1-81 <SAN>
C;Superfamily: saposin; saposin repeat homology
F;1-81/Domain: saposin repeat homology <SAP>

Query Match 1.7%; Score 9; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 357 CQEVDTYG 365

Db 47 CQEVDTYG 55
|||||

RESULT 7

AD1758
hypothetical protein lin2609 [imported] - Listeria innocua (strain Clp11262)

C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C;Accession: AD1758
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001

A;Authors: Krest, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat

ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AD1758

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-162 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC97836.1; PID:gi16415146; GSPDB:GN00178

A;Experimental source: strain Clp11262

C;Genetics:

A;Gene: lin2609

Query Match 1.5%; Score 8; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LFLASLL 11
|||||

Db 4 LFLASLL 11

RESULT 8

T47133

hypothetical protein DKFZp761P19121.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C;Accession: T47133

R;Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.

submitted to the Protein Sequence Database, March 2000

A;Reference number: Z24374

A;Accession: T47133

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-177 <AAA>

A;Cross-references: EMBL:ALJ62013

A;Experimental source: adult amygdala; clone DKFZp761P19121

C;Genetics:

A;Note: DKFZp761P19121.1

Query Match 1.5%; Score 8; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 EILAALEK 438
|||||

Db 146 EILAALEK 153

RESULT 9

HHFF23

heat shock protein 23 - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 16-Feb-1997

C;Accession: A02919

R;Ingolia, T.D.; Craig, E.A.

Proc. Natl. Acad. Sci. U.S.A. 79, 2360-2364, 1982

A;Title: Four small Drosophila heat shock proteins are related to each other and to mammi

A;Reference number: A93909; MUID:82248004; PMID:6285380


```

Qy      10  LLGAALAG 17
      |||||
Db      7  LLGAALAG 14

RESULT 12
S01929
fimbrial protein precursor - Bordetella pertussis
C:Species: Bordetella pertussis
C:Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 26-Aug-1999
C:Accession: S01929
R:Pedroni, P.; Riboli, B.; de Ferra, F.; Grandi, G.; Toma, S.; Arico, B.; Rappuoli, M.; Macrobio, L. 2, 539-543, 1988
A:Title: Cloning of a novel pilin-like gene from Bordetella pertussis: homology
A:Reference number: S01929; MUID:89013896; PMID:2902506
A:Accession: S01929
A:Molecule type: DNA
A:Residues: 1-201 <PED>
A:Cross-references: EMBL:Y00556; NID:g39750; PIDN:CAA68634.1; PID:g39751
C:Genetics:
A:Gene: fimX
C:Superfamily: type 1 fimbrial protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-201/Product: fimbrial protein #status predicted <MAT>

Query Match      1.5%; Score 8; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      10  LLGAALAG 17
      |||||
Db      7  LLGAALAG 14

RESULT 13
S56340
fimi protein - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 01-Mar-2002
C:Accession: S56540; E55245; E53066
R:Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 23, 2105-2119, 1995
A:Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region
A:Reference number: S56314; MUID:95334362; PMID:7610040
A:Accession: S56540
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-215 <BUR>
A:Cross-references: EMBL:U14003; NID:g1263172; PIDN:AAA97211.1; PID:g537156
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: E55245
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-215 <BLAT>
A:Cross-references: GB:AE000502; GB:U00096; NID:g2367374; PIDN:AAC77271.1; PID:
A:Experimental source: strain K-12, substrain MG1655
R:Marc, D.; Dho-Moulin, M.
submitted to the EMBL Data Library, September 1994
A:Description: Hypervariable sites within the major subunit of type 1 fimbriae,
A:Reference number: S53063
A:Accession: S53066
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 37-77, 'A', '79-107', 'S', '109-215 <MAR>
A:Cross-references: EMBL:Z37500; NID:g732680; PIDN:CAA85728.1; PID:g732684
C:Genetics:
A:Gene: fimi
A:Start codon: GTG

```

C;Superfamily: type 1 fimbrial protein

Query Match 1.5%; Score 8; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LFLLASLL 11
|||||||
Db 42 LFLLASLL 49

RESULT 14

E86129
fimbrial protein [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
C;Accession: E86129
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E86129
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-215 <STO>
A;Cross-references: GB:AE005174; NID:G12519322; PIDN:AAG59497.1; GSPDB:GN00145; UWGP:Z59
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: fimI
C;Superfamily: type 1 fimbrial protein

Query Match 1.5%; Score 8; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LFLLASLL 11
|||||||
Db 42 LFLLASLL 49

RESULT 15

B91288
fimbrial protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 02-Aug-2002
C;Accession: B91288
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B91288
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-215 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA838697.1; PID:gl33364752; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: EC85274
C;Superfamily: type 1 fimbrial protein

Query Match 1.5%; Score 8; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LFLLASLL 11
|||||||
Db 42 LFLLASLL 49

Search completed: May 5, 2004, 13:48:28
Job time : 16.595 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:33:49 ; Search time 42.2418 Seconds
(without alignments)

3438.471 Million cell updates/sec

Title: US-09-743-684A-1

Perfect score: 524

Sequence: 1 MYALFLIASLLGAALAGPVL.....NTETAACQNAVERCKRHVWN 524

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1138120 seqs, 277189581 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	524	100.0	524	9	US-09-870-759-60
2	524	100.0	524	10	US-09-751-708A-60
3	524	100.0	524	12	US-10-267-502-386
4	284	54.2	527	9	US-09-870-759-61
5	284	54.2	527	10	US-09-751-708A-61
6	284	54.2	527	14	US-10-060-036-73
7	260	49.6	479	10	US-09-978-418-40
8	209	39.9	209	14	US-10-043-487-340
9	193	36.8	523	9	US-09-767-007A-2
10	80	15.3	80	9	US-09-767-007A-3
11	80	15.3	80	9	US-09-753-126-3
12	80	15.3	80	15	US-10-330-697-3
13	80	15.3	592	9	US-09-753-126-4
14	80	15.3	592	15	US-10-330-697-4
15	40	7.6	40	10	US-09-780-438A-1

16	38	7.3	38	10	US-09-780-438A-2	Sequence 2, Appli
17	24	4.6	156	12	US-09-925-298-644	Sequence 644, App
18	156	4.6	156	14	US-10-102-806-644	Sequence 644, App
19	22	4.2	22	9	US-09-767-007A-1	Sequence 1, Appli
20	22	4.2	22	9	US-09-767-007A-11	Sequence 11, Appli
21	22	4.2	22	9	US-09-957-143-1	Sequence 1, Appli
22	20	3.8	554	14	US-10-205-194-176	Sequence 176, App
23	17	3.2	18	9	US-09-767-007A-5	Sequence 5, Appli
24	17	3.2	18	9	US-09-957-143-2	Sequence 2, Appli
25	15	2.9	15	9	US-09-802-617-2	Sequence 2, Appli
26	15	2.9	15	14	US-10-293-819-2	Sequence 2, Appli
27	12	2.3	12	9	US-09-802-617-1	Sequence 1, Appli
28	12	2.3	12	9	US-09-957-143-3	Sequence 3, Appli
29	12	2.3	12	14	US-10-293-819-1	Sequence 1, Appli
30	11	2.1	11	10	US-09-791-393-182	Sequence 182, App
31	11	2.1	11	10	US-09-791-389-182	Sequence 182, App
32	11	2.1	11	14	US-10-281-478-126	Sequence 126, App
33	10	1.9	15	14	US-10-281-478-74	Sequence 74, Appli
34	10	1.9	20	14	US-10-281-478-77	Sequence 77, Appli
35	9	1.7	14	14	US-10-281-478-75	Sequence 75, Appli
36	9	1.7	17	14	US-10-281-478-129	Sequence 129, App
37	8	1.5	215	9	US-09-912-020-362	Sequence 362, App
38	8	1.5	269	12	US-10-243-552-937	Sequence 937, App
39	8	1.5	286	15	US-10-108-260A-4623	Sequence 4623, Ap
40	8	1.5	354	12	US-10-424-599-279814	Sequence 279814,
41	8	1.5	361	12	US-10-425-114-46329	Sequence 46329, A
42	8	1.5	363	12	US-10-425-114-72368	Sequence 72368, A
43	8	1.5	391	14	US-10-017-161-2324	Sequence 2324, Ap
44	8	1.5	391	15	US-10-292-798-1970	Sequence 1970, Ap
45	8	1.5	521	12	US-10-412-699B-82	Sequence 82, Appli

ALIGNMENTS

RESULT 1

US-09-870-759-60
; Sequence 60, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-60

Query Match	100.0%;	Score 524;	DB 9;	Length 524;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 524;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MYALFLIASLLGAALAGPVLGLKECTRGSAVWCNVTASDCGAVKHCLQTVWNKPTVKS	60	
Db	1	MYALFLIASLLGAALAGPVLGLKECTRGSAVWCNVTASDCGAVKHCLQTVWNKPTVKS	60	
Qy	61	LPDCICKDVVTAAGDMLKDNATEEILVLYEKTCDMLPKNMSASCKEIVDSYLPVILDI	120	
Db	61	LPDCICKDVVTAAGDMLKDNATEEILVLYEKTCDMLPKNMSASCKEIVDSYLPVILDI	120	
Qy	121	IKGEMSRPGVCSALNLCESLQKHLAEINLHOKLESNKIPELDMTEVVAPEMANIPILLY	180	
Db	121	IKGEMSRPGVCSALNLCESLQKHLAEINLHOKLESNKIPELDMTEVVAPEMANIPILLY	180	
Qy	181	PQDGRSKPQKONGDVVCQDCIQMTDITQAVRTNFTFQALVEHVKECDRLGPGMADI	240	

Db 181 PDGPRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Qy 241 CKNYISQYSEIATOMMMHMQPKIEICALVGFCDEVKEMPMQTLVPAKASKNVI PALELVE 300
Db 241 CKNYISQYSEIATOMMMHMQPKIEICALVGFCDEVKEMPMQTLVPAKASKNVI PALELVE 300
Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPSLSEECQEV 360
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPSLSEECQEV 360
Qy 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVHTQPKDGGFCEVCKKLVGYLD 420
Db 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVHTQPKDGGFCEVCKKLVGYLD 420
Qy 421 RNLEKNTKQEIILAALEKSCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNTKQEIILAALEKSCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVWN 524

RESULT 2

US-09-751-708A-60
; Sequence 60, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; CURRENT APPLICATION NUMBER: US/09/751.708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-751-708A-60

Query Match 100.0%; Score 524; DB 10; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKNKPTVKS 60
Db 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKNKPTVKS 60
Qy 61 LPCDICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGEMSPGVCESALNLCESLQKHLAEHLNKHOKLESNKIPELDMTEVAPFMANIPLLLY 180
Db 121 IKGEMSPGVCESALNLCESLQKHLAEHLNKHOKLESNKIPELDMTEVAPFMANIPLLLY 180
Qy 181 PDGPRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Db 181 PDGPRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Qy 241 CKNYISQYSEIATOMMMHMQPKIEICALVGFCDEVKEMPMQTLVPAKASKNVI PALELVE 300
Db 241 CKNYISQYSEIATOMMMHMQPKIEICALVGFCDEVKEMPMQTLVPAKASKNVI PALELVE 300
Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPSLSEECQEV 360
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPSLSEECQEV 360
Qy 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVHTQPKDGGFCEVCKKLVGYLD 420

Db 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVHTQPKDGGFCEVCKKLVGYLD 420
Qy 421 RNLEKNTKQEIILAALEKSCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNTKQEIILAALEKSCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVWN 524

RESULT 3

US-10-267-502-386
; Sequence 386, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 386
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-267-502-386

Query Match 100.0%; Score 524; DB 12; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKNKPTVKS 60
Db 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKNKPTVKS 60
Qy 61 LPCDICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKDVVTAAGDMLKDNATEBEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGEMSPGVCESALNLCESLQKHLAEHLNKHOKLESNKIPELDMTEVAPFMANIPLLLY 180
Db 121 IKGEMSPGVCESALNLCESLQKHLAEHLNKHOKLESNKIPELDMTEVAPFMANIPLLLY 180
Qy 181 PDGPRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Db 181 PDGPRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Qy 241 CKNYISQYSEIATOMMMHMQPKIEICALVGFCDEVKEMPMQTLVPAKASKNVI PALELVE 300
Db 241 CKNYISQYSEIATOMMMHMQPKIEICALVGFCDEVKEMPMQTLVPAKASKNVI PALELVE 300
Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPSLSEECQEV 360
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPSLSEECQEV 360
Qy 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVHTQPKDGGFCEVCKKLVGYLD 420
Db 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVHTQPKDGGFCEVCKKLVGYLD 420
Qy 421 RNLEKNTKQEIILAALEKSCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNTKQEIILAALEKSCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVWN 524

RESULT 4

US-09-870-759-61

```
; Sequence 61, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 61
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-870-759-61

Query Match      54.2%; Score 284; DB 9; Length 527;
Best Local Similarity 99.4%; Pred. No. 1.1e-267;
Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 MYALFLLASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWKNKPTVKS 60
Db 1 MYALFLLASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWKNKPTVKS 60
Qy 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVYVAPFMANIPLLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVYVAPFMANIPLLY 180
Qy 181 PODGPRSKPOPKONGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Db 181 PODGPRSKPOPKONGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Qy 241 CKNYISOYSEIAIOMMHM--OPKEICALVGFCDVEKMPQTLVPAKASKNVIPALE 297
Db 241 CKNYISOYSEIAIOMMHMDDQPKETCALVGFCDVEKMPQTLVPAKASKNVIPALE 300
Qy 298 LVEPIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFRCMSKLPKSLSEEC 357
Db 301 LVEPIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFRCMSKLPKSLSEEC 360
Qy 358 QEVVDYTGSSILSLLEEVSPVLCVSMHLCSGTRLPALTVHVTQPKDGGFCEVCKKLVG 417
Db 361 QEVVDYTGSSILSLLEEVSPVLCVSMHLCSGTRLPALTVHVTQPKDGGFCEVCKKLVG 420
Qy 418 YLDRNLEKNSKTKQELIAALEKGCFLPDPYQKQCDQFVAEYEPVLIILVEVMDPSFVCL 477
Db 421 YLDRNLEKNSKTKQELIAALEKGCFLPDPYQKQCDQFVAEYEPVLIILVEVMDPSFVCL 480
Qy 478 KIGACPSAHKPLLGTEKRCINGSPSYWCQNTETAACQNAVEHCKRHVWN 524
Db 481 KIGACPSAHKPLLGTEKRCINGSPSYWCQNTETAACQNAVEHCKRHVWN 527

RESULT 5
US-09-751-708A-61
; Sequence 61, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: Patentin version 3.1
```

```
; SEQ ID NO 61
; LENGTH: 527
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-751-708A-61

Query Match      54.2%; Score 284; DB 10; Length 527;
Best Local Similarity 99.4%; Pred. No. 1.1e-267;
Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 MYALFLLASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWKNKPTVKS 60
Db 1 MYALFLLASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWKNKPTVKS 60
Qy 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVYVAPFMANIPLLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVYVAPFMANIPLLY 180
Qy 181 PODGPRSKPOPKONGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Db 181 PODGPRSKPOPKONGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Qy 241 CKNYISOYSEIAIOMMHM--OPKEICALVGFCDVEKMPQTLVPAKASKNVIPALE 297
Db 241 CKNYISOYSEIAIOMMHMDDQPKETCALVGFCDVEKMPQTLVPAKASKNVIPALE 300
Qy 298 LVEPIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFRCMSKLPKSLSEEC 357
Db 301 LVEPIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFRCMSKLPKSLSEEC 360
Qy 358 QEVVDYTGSSILSLLEEVSPVLCVSMHLCSGTRLPALTVHVTQPKDGGFCEVCKKLVG 417
Db 361 QEVVDYTGSSILSLLEEVSPVLCVSMHLCSGTRLPALTVHVTQPKDGGFCEVCKKLVG 420
Qy 418 YLDRNLEKNSKTKQELIAALEKGCFLPDPYQKQCDQFVAEYEPVLIILVEVMDPSFVCL 477
Db 421 YLDRNLEKNSKTKQELIAALEKGCFLPDPYQKQCDQFVAEYEPVLIILVEVMDPSFVCL 480
Qy 478 KIGACPSAHKPLLGTEKRCINGSPSYWCQNTETAACQNAVEHCKRHVWN 524
Db 481 KIGACPSAHKPLLGTEKRCINGSPSYWCQNTETAACQNAVEHCKRHVWN 527

RESULT 6
US-10-060-036-73
; Sequence 73, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 527
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-060-036-73

Query Match      54.2%; Score 284; DB 14; Length 527;
Best Local Similarity 99.4%; Pred. No. 1.1e-267;
```

Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60

QY 61 LPCDICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120

QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWVAFPMANIPLLLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWVAFPMANIPLLLY 180

QY 181 PQDGRSKPQKNDGVDVDCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Db 181 PQDGRSKPQKNDGVDVDCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240

QY 241 CKNYISQYSEIAIQMMHMQ 260
Db 241 CKNYISQYSEIAIQMMHMQ 260

RESULT 8
US-10-043-487-340
; Sequence 340, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEBRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 340
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Shigella Flexneri
US-10-043-487-340

Query Match 39.9%; Score 209; DB 14; Length 209;
Best Local Similarity 100.0%; Pred. No. 5.7e-195;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 OKLESNKIPELDMTEVWVAFPMANIPLLYPQDGRSKPQKNDGVDVDCDCIQMTVDIQT 210
Db 1 OKLESNKIPELDMTEVWVAFPMANIPLLYPQDGRSKPQKNDGVDVDCDCIQMTVDIQT 60

QY 211 AVRTNSTFVQALVEHVKEECDRLGPGMADICKNYISQYSEIAIQMMHMQKEICLVGF 270
Db 61 AVRTNSTFVQALVEHVKEECDRLGPGMADICKNYISQYSEIAIQMMHMQKEICLVGF 120

QY 271 CDEVKEMPQTLVPAKVASKNVIPALELVEPIKKEHVPKASDVYCEVCEFLVKEVTKLID 330
Db 121 CDEVKEMPQTLVPAKVASKNVIPALELVEPIKKEHVPKASDVYCEVCEFLVKEVTKLID 180

QY 331 NNKTEKILDAFDKMSKLPKSLSEECQE 359
Db 181 NNKTEKILDAFDKMSKLPKSLSEECQE 209

RESULT 9
US-09-767-007A-2
; Sequence 2, Application US/09767007A
; Patent No. US20020077275A1
; GENERAL INFORMATION:
; APPLICANT: John S. O'Brien
; APPLICANT: Yasuo Kishimoto
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
; FILE REFERENCE: SAPOSIN C AND NEUROTROPIC PEPTIDES DERIVED THEREFROM
; CURRENT APPLICATION NUMBER: US/09/767,007A
; CURRENT FILING DATE: 2001-01-22

Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60

QY 61 LPCDICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120

QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWVAFPMANIPLLLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWVAFPMANIPLLLY 180

QY 181 PQDGRSKPQKNDGVDVDCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Db 181 PQDGRSKPQKNDGVDVDCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240

QY 241 CKNYISQYSEIAIQMMHMQ 260
Db 241 CKNYISQYSEIAIQMMHMQ 260

QY 298 LVEPIKKEHVPKASDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEEC 357
Db 301 LVEPIKKEHVPKASDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEEC 360

QY 358 QEVVDYTGSSILSILLESVPELVCSMLHLCSTGRTPALTVHVTQPKDGGCEVCKKLVG 417
Db 361 QEVVDYTGSSILSILLESVPELVCSMLHLCSTGRTPALTVHVTQPKDGGCEVCKKLVG 420

QY 418 YLDNRLEKNSTKQETILALEKCSFLPDYQKQCDQFVAEYEPVLIILVEVMDPSFVCL 477
Db 421 YLDNRLEKNSTKQETILALEKCSFLPDYQKQCDQFVAEYEPVLIILVEVMDPSFVCL 480

QY 478 KIGACPSAHKPLLTGTEKICWGPSYWCNTETAACNAVEHCKRHVN 524
Db 481 KIGACPSAHKPLLTGTEKICWGPSYWCNTETAACNAVEHCKRHVN 527

RESULT 7
US-09-978-418-40
; Sequence 40, Application US/09978418
; Publication No. US20030118997A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephan
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 142.US5.REG
; CURRENT APPLICATION NUMBER: US/09/978,418
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/311,305
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/314,734
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/318,204
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/326,470
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: JPatent
; SEQ ID NO 40
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-418-40

Query Match 49.6%; Score 260; DB 10; Length 479;
Best Local Similarity 100.0%; Pred. No. 2.5e-244;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60

```
; PRIOR APPLICATION NUMBER: 08/958,970
; PRIOR FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 523
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-767-007A-2

Query Match          36.8%; Score 193; DB 9; Length 523;
Best Local Similarity 100.0%; Pred. No. 5.1e-179; Indels 0; Gaps 0;
Matches 193; Conservative 0; Mismatches 0;

Qy 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTVWNKPTVKS 60
Db 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTVWNKPTVKS 60

Qy 61 LPDCICKDVVTAAGDMLKDNATBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPDCICKDVVTAAGDMLKDNATBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120

Qy 121 IKGEMSPGVCNALNCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPILLY 180
Db 121 IKGEMSPGVCNALNCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPILLY 180

Qy 181 PQDGPSPKPOPKD 193
Db 181 PQDGPSPKPOPKD 193

RESULT 10
US-09-767-007A-3
; Sequence 3, Application US/09767007A
; Patent No. US2002007275A1
; GENERAL INFORMATION:
; APPLICANT: John S. O'Brien
; APPLICANT: Yasuo Kishimoto
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
; FILE OF INVENTION: SAPOSIN C AND NEUROTROPIC PEPTIDES DERIVED THEREFROM
; FILE REFERENCE: MYELOS.DC1C1
; CURRENT APPLICATION NUMBER: US/09/767,007A
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 08/958,970
; PRIOR FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide
US-09-767-007A-3

Query Match          15.3%; Score 80; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.2e-69; Indels 0; Gaps 0;
Matches 80; Conservative 0; Mismatches 0;

Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 370
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 60

Qy 371 ILLEEVSPELVCSMLHLCSG 390
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 11
US-09-753-126-3
; Sequence 3, Application US/09753126
; Patent No. US20020127219A1
; GENERAL INFORMATION:
```

```
; APPLICANT: OKKELS, JENS SIGURD
; APPLICANT: JENSEN, ANNE DAM
; APPLICANT: HALKIER, TORBEN
; APPLICANT: JENSEN, RIKKE BOLDING
; TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME
; TITLE OF INVENTION: ACTIVATORS
; FILE REFERENCE: 31-0006000S
; CURRENT APPLICATION NUMBER: US/09/753,126
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PA 1999 01891
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 60/174,652
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PA 200 00865
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/210,984
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 60/211,124
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: PA 2000 01027
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/217,497
; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-753-126-3

Query Match          15.3%; Score 80; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.2e-69; Indels 0; Gaps 0;
Matches 80; Conservative 0; Mismatches 0;

Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 370
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 60

Qy 371 ILLEEVSPELVCSMLHLCSG 390
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 12
US-10-330-697-3
; Sequence 3, Application US/10330697
; Publication No. US20040009165A1
; GENERAL INFORMATION:
; APPLICANT: OKKELS, JENS SIGURD
; APPLICANT: JENSEN, ANNE DAM
; APPLICANT: HALKIER, TORBEN
; APPLICANT: JENSEN, RIKKE BOLDING
; TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME
; TITLE OF INVENTION: ACTIVATORS
; FILE REFERENCE: 31-0006000S
; CURRENT APPLICATION NUMBER: US/10/330,697
; CURRENT FILING DATE: 2002-12-27
; PRIOR APPLICATION NUMBER: US/09/753,126
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PA 1999 01891
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 60/174,652
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PA 200 00865
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/210,984
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 60/211,124
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: PA 2000 01027
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/217,497
```

```

; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-330-697-3

```

Query Match	15.3%	Score 80;	DB 15;	Length 80;
Beat Local Similarity	100.0%;	Pred. No. 1.2e-69;		
Matches 80;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	311	SDVCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSECCOEVDVTGSSILS	370	
Db	1	SDVCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSECCOEVDVTGSSILS	60	
QY	371	ILLEVPSPVCSMLHLCSG	390	
Db	61	ILLEVPSPVCSMLHLCSG	80	

RESULT 13
 US-09-753-126-4
 ; Sequence 4, Application US/09753126
 ; Patent No. US2002012719A1
 ; GENERAL INFORMATION:
 APPLICANT: OKKELS, JENS SIGURD
 APPLICANT: JENSEN, ANNE DAM
 APPLICANT: HALKIER, TORBEN
 APPLICANT: JENSEN, RIKKE BOLDING
 TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME
 TITLE OF INVENTION: ACTIVATORS
 FILE REFERENCE: 31-000600US
 CURRENT APPLICATION NUMBER: US/09/753,126
 CURRENT FILING DATE: 2001-06-11
 PRIOR APPLICATION NUMBER: PA 1999 01891
 PRIOR FILING DATE: 1999-12-30
 PRIOR APPLICATION NUMBER: 60/174,652
 PRIOR FILING DATE: 2000-01-06
 PRIOR APPLICATION NUMBER: PA 200 00865
 PRIOR FILING DATE: 2000-06-02
 PRIOR APPLICATION NUMBER: 60/210,984
 PRIOR FILING DATE: 2000-06-12
 PRIOR APPLICATION NUMBER: 60/211,124
 PRIOR FILING DATE: 2000-06-12
 PRIOR APPLICATION NUMBER: PA 2000 01027
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: 60/217,497
 PRIOR FILING DATE: 2000-07-11
 NUMBER OF SEQ ID NOS: 147
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 4
 LENGTH: 592
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Chimeric
 OTHER INFORMATION: SapC-linker-GCB polypeptide
 S-09-753-126-4

```

RESULT 14
US-10-330-697-4
; Sequence 4, Application US/10330697
; Publication No. US20040009165A1
; GENERAL INFORMATION:
; APPLICANT: OKKELS, JENS SIGURD
; APPLICANT: JENSEN, ANNE DAM
; APPLICANT: HALKIER, TORBEN
; APPLICANT: JENSEN, RIKKE BOLDING
; TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME
; TITLE OF INVENTION: ACTIVATORS
; FILE REFERENCE: 31-000600US
; CURRENT APPLICATION NUMBER: US/10/330,697
; CURRENT FILING DATE: 2002-12-27
; PRIOR APPLICATION NUMBER: US/09/753,126
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PA 1999 01891
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 60/174,652
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PA 200 00865
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/210,984
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 60/211,124
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: PA 2000 01027
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/217,497
; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 592
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; OTHER INFORMATION: SapC-linker-OCB polypeptide
; US-10-330-697-4

```

```

RESULT 15
US-09-780-438A-1
; Sequence 1, Application US/09780438A
; Publication No. US20030095999A1
; GENERAL INFORMATION:
; APPLICANT: Cincinnati Childrens Hospital Research Foundation
; TITLE OF INVENTION: Fungogenic Properties of Saposin C and Related Proteins and Polypeptides
; TITLE OF INVENTION: fox Application to Transmembrane Drug Delivery Systems
; FILE REFERENCE: 10872/0474352
; CURRENT APPLICATION NUMBER: US/09/780,438A
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: US 60/181,754
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 40
; TYPE: PRT

```


; ORGANISM: Homo sapiens
US-09-780-438A-1

US-09-780-438A-1

Query Match 7.6%; Score 40; DB 10; Length 40;

Best Local Similarity 100.0%; Pred. No. 5.7e-31;

Best local similarity 100.0%, freq. NO. 3.7E-31;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCSKLP 350

1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDAPDKMCSKL.P 40

Search completed: May 5, 2004, 13:43:39
Job time : 43.2418 secs

Job time : 43.2418 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:28:28 ; Search time 16.595 Seconds
(without alignments)
1630.129 Million cell updates/sec

Title: US-09-743-684A-1
Perfect score: 524
Sequence: 1 MYALFLIASLGAALAGPVL.....NTETAQCNAVERKGRHVN 524

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTCUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	524	100.0	524	4	US-09-352-548-1
2	224	42.7	523	1	US-08-100-247-2
3	224	42.7	523	1	US-08-232-513A-3
4	224	42.7	523	4	US-08-756-031-2
5	193	36.8	523	1	US-08-483-146A-2
6	193	36.8	523	1	US-08-484-594A-2
7	193	36.8	523	4	US-09-076-258A-2
8	81	15.5	81	4	US-09-352-548-2
9	80	15.3	80	1	US-08-100-247-3
10	80	15.3	80	1	US-08-483-146A-3
11	80	15.3	80	1	US-08-232-513A-4
12	80	15.3	80	1	US-08-484-594A-3
13	80	15.3	80	4	US-09-076-258A-3
14	80	15.3	80	4	US-08-756-031-3
15	75	14.3	80	2	US-08-584-671-15
16	75	14.3	80	3	US-09-027-376-15
17	69	13.2	80	3	US-09-094-192-15
18	61	11.6	61	2	US-08-584-671-13
19	61	11.6	61	3	US-09-027-376-13
20	61	11.6	61	3	US-09-094-192-13
21	61	11.6	69	3	US-09-268-070-2
22	22	4.2	22	1	US-08-100-247-1
23	22	4.2	22	1	US-08-483-146A-1
24	22	4.2	22	1	US-08-483-146A-11
25	22	4.2	22	1	US-08-232-513A-1
26	22	4.2	22	1	US-08-232-513A-20
27	22	4.2	22	1	US-08-484-594A-1

28	22	4.2	22	1	US-08-484-594A-11	Sequence 11, Appl
29	22	4.2	22	3	US-09-231-159-1	Sequence 1, Appl
30	22	4.2	22	3	US-08-611-307-1	Sequence 1, Appl
31	22	4.2	22	4	US-09-148-030-1	Sequence 1, Appl
32	22	4.2	22	4	US-09-076-258A-1	Sequence 1, Appl
33	22	4.2	22	4	US-08-756-031-1	Sequence 1, Appl
34	22	4.2	22	4	US-08-756-031-1	Sequence 1, Appl
35	18	3.4	18	4	US-09-352-548-17	Sequence 17, Appl
36	17	3.2	17	4	US-09-352-548-14	Sequence 14, Appl
37	17	3.2	17	4	US-09-352-548-15	Sequence 15, Appl
38	17	3.2	18	1	US-08-100-247-5	Sequence 5, Appl
39	17	3.2	18	1	US-08-483-146A-5	Sequence 5, Appl
40	17	3.2	18	1	US-08-232-513A-6	Sequence 6, Appl
41	17	3.2	18	1	US-08-484-594A-5	Sequence 5, Appl
42	17	3.2	18	3	US-09-231-159-20	Sequence 20, Appl
43	17	3.2	18	3	US-08-611-307-20	Sequence 20, Appl
44	17	3.2	18	4	US-09-148-030-2	Sequence 2, Appl
45	17	3.2	18	4	US-09-076-258A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-352-548-1
; Sequence 1, Application US/09352548
; Patent No. 6500431
; GENERAL INFORMATION:
; APPLICANT: Gill, Parkash S.
; APPLICANT: Parkash S. Gill, M.D., Inc.
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
; FILE REFERENCE: 017986-000410US
; CURRENT APPLICATION NUMBER: US/09/352,548
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: US 60/092,647
; EARLIER FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: proasapoin
; NAME/KEY: PEPTIDE
; LOCATION: (195)..(275)
; OTHER INFORMATION: Saposin B
; US-09-352-548-1

Query Match	100.0%;	Score 524;	DB 4;	Length 524;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 524;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MYALFLIASLGAALAGPVLGLKECTRGSAVWCONVTASDCGAVKHCLQTVWNKPTVKS	60	
Db	1	MYALFLIASLGAALAGPVLGLKECTRGSAVWCONVTASDCGAVKHCLQTVWNKPTVKS	60	
Qy	61	LPDCICKDVVTAAGDMLKDNATREELVYLEKTCMDLWPKNMSASCKEIVDSYLPVILDI	120	
Db	61	LPDCICKDVVTAAGDMLKDNATREELVYLEKTCMDLWPKNMSASCKEIVDSYLPVILDI	120	
Qy	121	IKGEMSRPGVCSALNLCESLQKHAEINHKQLESNKIPELDMTEVVAPPMANIPILLY	180	
Db	121	IKGEMSRPGVCSALNLCESLQKHAEINHKQLESNKIPELDMTEVVAPPMANIPILLY	180	
Qy	181	PQDGRSKPQKNDGVDVCDCIQMVTDTQTAVRTNSTFFQALVEHVKECDRLGFGMADI	240	
Db	181	PQDGRSKPQKNDGVDVCDCIQMVTDTQTAVRTNSTFFQALVEHVKECDRLGFGMADI	240	
Qy	241	CKNYISYSEIAIQMMHMQPKKEICALVGFCDEKVENPMQTLVPAKASKNVIIPALELVE	300	
Db	241	CKNYISYSEIAIQMMHMQPKKEICALVGFCDEKVENPMQTLVPAKASKNVIIPALELVE	300	

QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQV 360
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQV 360
QY 361 VDTYGSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420
Db 361 VDTYGSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420
QY 421 RNLEKNSTKQEIILAALEKGSFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTKQEIILAALEKGSFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
QY 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHWN 524

RESULT 2

US-08-100-247-2
; Sequence 2, Application US/08100247
; Patent No. 5571787
; GENERAL INFORMATION:
; APPLICANT: O'BRIEN, JOHN S.
; APPLICANT: KISHIMOTO, YASUO
; TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/100,247
; FILING DATE: 19930730
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: OBRIEN.002A
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; IMMEDIATE SOURCE:
; CLONE: PROSAPOSIN
US-08-100-247-2

Query Match 42.7%; Score 224; DB 1; Length 523;
Best Local Similarity 100.0%; Pred. No. 3.6e-207;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQV 360
Db 300 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQV 359
QY 361 VDTYGSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420

Db 360 VDTYGSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 419
QY 421 RNLEKNSTKQEIILAALEKGSFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Db 420 RNLEKNSTKQEIILAALEKGSFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 479
QY 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHWN 524
Db 480 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHWN 523

RESULT 3

US-08-232-513A-3
; Sequence 3, Application US/08232513A
; Patent No. 5700909
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides
; TITLE OF INVENTION: as Therapeutic Agents
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,513A
; FILING DATE: 21-APR-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/100,247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 1643
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..523
; OTHER INFORMATION: /label= Hum_prosaposin
US-08-232-513A-3

Query Match 42.7%; Score 224; DB 1; Length 523;
Best Local Similarity 100.0%; Pred. No. 3.6e-207;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQV 360
Db 300 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQV 359
QY 361 VDTYGSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420
Db 360 VDTYGSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 419
QY 421 RNLEKNSTKQEIILAALEKGSFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480

Db 420 RNLEKSTKOEIIAALAEKGCFLDPYQKQCDQFVAEYEPVLIEILVEMDPSFVCLKIG 479

Qy 481 ACPHAKPLLGTEKCIWGPSPYWCNTETAAQCNVAHECHKEHVN 524

Db 480 ACPHAKPLLGTEKCIWGPSPYWCNTETAAQCNVAHECHKEHVN 523

RESULT 4

US-08-756-031-2

; Sequence 2, Application US/08756031

; Patent No. 6590074

; GENERAL INFORMATION:

; APPLICANT: O'BRIEN, JOHN S.

; APPLICANT: KISHIMOTO, YASUO

; TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR

; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR

; CITY: NEWPORT BEACH

; STATE: CA

; COUNTRY: USA

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/756,031

; FILING DATE: 26-NOV-1996

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/100,247

; FILING DATE: 30-JUL-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Israelsen, Ned A.

; REGISTRATION NUMBER: 29,655

; REFERENCE/DOCKET NUMBER: O'BRIEN.002A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-235-8550

; TELEFAX: 619-235-0176

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 523 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

; IMMEDIATE SOURCE:

; CLONE: PROSAPOSIN

US-08-756-031-2

Query Match 42.7%; Score 224; DB 4; Length 523;

Best Local Similarity 100.0%; Pred. No. 3.6e-207;

Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 301 PIKHEVPKSDVYCEVFLVKEVTKLIDNNKTEKEILDAFDKWCCKLPSLSEECQEV 360

Db 300 PIKHEVPKSDVYCEVFLVKEVTKLIDNNKTEKEILDAFDKWCCKLPSLSEECQEV 359

Qy 361 VDTGSSILSILLLEVPVCSMLHLCSTGRLPALTVHTVTPKDGCFCEVCKLVGYLD 420

Db 360 VDTGSSILSILLLEVPVCSMLHLCSTGRLPALTVHTVTPKDGCFCEVCKLVGYLD 419

Qy 421 RNLEKSTKOEIIAALAEKGCFLDPYQKQCDQFVAEYEPVLIEILVEMDPSFVCLKIG 480

Db 420 RNLEKSTKOEIIAALAEKGCFLDPYQKQCDQFVAEYEPVLIEILVEMDPSFVCLKIG 479

Qy 481 ACPHAKPLLGTEKCIWGPSPYWCNTETAAQCNVAHECHKEHVN 524

Db 480 ACPHAKPLLGTEKCIWGPSPYWCNTETAAQCNVAHECHKEHVN 523

RESULT 5

US-08-483-146A-2

; Sequence 2, Application US/08483146A

; Patent No. 5696080

; GENERAL INFORMATION:

; APPLICANT: O'Brien, John S.

; APPLICANT: Kishimoto, Yasuo

; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS

; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED THEREFROM

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson and Bear

; STREET: 620 Newport Center Blvd. 16th Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: USA

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/483,146A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Israelsen, Ned A.

; REGISTRATION NUMBER: 29,655

; REFERENCE/DOCKET NUMBER: MYELOS.002DV1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-235-8550

; TELEFAX: 619-235-0176

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 523 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: N-terminal

US-08-483-146A-2

Query Match 36.8%; Score 193; DB 1; Length 523;

Best Local Similarity 100.0%; Pred. No. 2.5e-177;

Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALFLILASLLGALAGPVLGLKECTRGSAVWCNVKTASDCGAVKHCLQTVWKNKPTVKS 60

Db 1 MYALFLILASLLGALAGPVLGLKECTRGSAVWCNVKTASDCGAVKHCLQTVWKNKPTVKS 60

Qy 61 LPCDICKQVVVTAAGDMLKDNATBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120

Db 61 LPCDICKQVVVTAAGDMLKDNATBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120

Qy 121 IKGEMSRPGVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVVVAFPMANIPILLY 180

Db 121 IKGEMSRPGVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVVVAFPMANIPILLY 180

Qy 181 PQDGRSKPQPKD 193

Db 181 PQDGRSKPQPKD 193

RESULT 6

US-08-484-594A-2
; Sequence 2, Application US/08484594A
; Patent No. 5714459
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; APPLICANT: Kishimoto, Yasuo
; TITLE OF INVENTION: USE OF PROSAPIN AND NEUROTROPHIC PEPTIDES
; TITLE OF INVENTION: DERIVED THEREFROM
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,594A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/100,247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: MYELOS.002DV2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-484-594A-2

Query Match 36.8%; Score 193; DB 1; Length 523;
Best Local Similarity 100.0%; Pred. No. 2.5e-177;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYALFLASLLGALAGPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWNKPTVKS 60
Db 1 MYALFLASLLGALAGPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWNKPTVKS 60
Qy 61 LPCDICKDVVTAAGDMLKDNATBEEILVLEKTCDLWLPKNMSASCKEIVDSYLPVLDI 120
Db 61 LPCDICKDVVTAAGDMLKDNATBEEILVLEKTCDLWLPKNMSASCKEIVDSYLPVLDI 120
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180
Qy 181 PQGPRSKPOPKD 193
Db 181 PQGPRSKPOPKD 193

RESULT 7

US-09-076-258A-2
; Sequence 2, Application US/09076258A
; Patent No. 6559124
; GENERAL INFORMATION:

; APPLICANT: O'Brien, John S.
; APPLICANT: Kishimoto, Yasuo
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
; TITLE OF INVENTION: COMPRISING PROSAPIN AND NEUROTROPHIC PEPTIDES DERIVED
; TITLE OF INVENTION: THEREFROM
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Blvd. 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/076,258A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/958,970
; FILING DATE: 28-OCT-97
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/483,146
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: 08/100,247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: MYELOS.2DV1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-09-076-258A-2

Query Match 36.8%; Score 193; DB 4; Length 523;
Best Local Similarity 100.0%; Pred. No. 2.5e-177;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYALFLASLLGALAGPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWNKPTVKS 60
Db 1 MYALFLASLLGALAGPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWNKPTVKS 60
Qy 61 LPCDICKDVVTAAGDMLKDNATBEEILVLEKTCDLWLPKNMSASCKEIVDSYLPVLDI 120
Db 61 LPCDICKDVVTAAGDMLKDNATBEEILVLEKTCDLWLPKNMSASCKEIVDSYLPVLDI 120
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180
Qy 181 PQGPRSKPOPKD 193
Db 181 PQGPRSKPOPKD 193

RESULT 8

US-09-352-548-2
; Sequence 2, Application US/09352548
; Patent No. 6500431

GENERAL INFORMATION:
APPLICANT: Gill, Parkash S.
APPLICANT: Parkash S. Gill, M.D., Inc.
TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
FILE REFERENCE: 017986-000410US
CURRENT APPLICATION NUMBER: US/09/352,548
CURRENT FILING DATE: 1999-07-12
EARLIER APPLICATION NUMBER: US 60/092,647
EARLIER FILING DATE: 1998-07-13
NUMBER OF SEQ ID NOS: 59
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 81
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Saposin B
US-09-352-548-2

Query Match 15.5%; Score 81; DB 4; Length 81;
Best Local Similarity 100.0%; Pred. No. 2.9e-70;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 195 GDVQCDCIQWTDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADICKNYSQYSEIAIQ 254
Db 1 GDVQCDCIQWTDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADICKNYSQYSEIAIQ 60
Qy 255 MMHMQPKEICALVGFCDDEVK 275
Db 61 MMHMQPKEICALVGFCDDEVK 81

RESULT 9
US-08-100-247-3
Sequence 3, Application US/08100247
Patent No. 5571787
GENERAL INFORMATION:
APPLICANT: O'BRIEN, JOHN S.
APPLICANT: KISHIMOTO, YASUO
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/100,247
FILING DATE: 19930730
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: OBRIEN.002A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO

FRAGMENT TYPE: internal
IMMEDIATE SOURCE:
CLONE: SAPOSIN C
US-08-100-247-3
Query Match 15.3%; Score 80; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.6e-69;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDYTGSSILS 370
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDYTGSSILS 60
Qy 371 ILLEEVSPELVCSMLHLCSG 390
Db 61 ILLEEVSPELVCSMLHLCSG 80
RESULT 10
US-08-483-146A-3
Sequence 3, Application US/08483146A
Patent No. 5696080
GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
APPLICANT: Kishimoto, Yasuo
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED
TITLE OF INVENTION: THEREFROM
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Blvd. 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/08/483,146A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.002DV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-483-146A-3
Query Match 15.3%; Score 80; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.6e-69;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDYTGSSILS 370
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDYTGSSILS 60

Qy 371 ILLEEVSPELVCSMLHLCSG 390
 Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 11

US-08-232-513A-4
 ; Sequence 4, Application US/08232513A
 ; Patent No. 5700909
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Brien, John S.
 ; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides
 ; TITLE OF INVENTION: as Therapeutic Agents
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Campbell & Flores LLP
 ; STREET: 4370 La Jolla Village Drive, Suite 700
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 92122
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/232,513A
 ; FILING DATE: 21-APR-1994
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/100,247
 ; FILING DATE: 30-JUL-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Campbell, Cathryn A.
 ; REGISTRATION NUMBER: 31,815
 ; REFERENCE/DOCKET NUMBER: P-UD 1643
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 535-9001
 ; TELEFAX: (619) 535-8949
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 80 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FEATURE:
 ; NAME/KEY: Peptide
 ; LOCATION: 1-80
 ; OTHER INFORMATION: /label= Saposin_C
 ; US-08-232-513A-4

Query Match 15.3%; Score 80; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 2.6e-69;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDYTGSSILS 370
 Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDYTGSSILS 60

Qy 371 ILLEEVSPELVCSMLHLCSG 390
 Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 12

US-08-484-594A-3
 ; Sequence 3, Application US/08484594A
 ; Patent No. 5714459
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Brien, John S.
 ; APPLICANT: Kishimoto, Yasuo

; TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUROTROPHIC PEPTIDES
 ; TITLE OF INVENTION: DERIVED THEREFROM
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Knobbe, Martens, Olson and Bear
 ; STREET: 620 Newport Center Drive, Sixteenth Floor
 ; CITY: Newport Beach
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92660
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/484,594A
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/100,247
 ; FILING DATE: 30-JUL-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Israel, Ned A.
 ; REGISTRATION NUMBER: 29,655
 ; REFERENCE/DOCKET NUMBER: MYELOS.002DV2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-235-8550
 ; TELEFAX: 619-235-0176
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 80 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FRAGMENT TYPE: internal
 ; US-08-484-594A-3

Query Match 15.3%; Score 80; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 2.6e-69;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDYTGSSILS 370
 Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDYTGSSILS 60

Qy 371 ILLEEVSPELVCSMLHLCSG 390
 Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 13

US-09-076-258A-3
 ; Sequence 3, Application US/09076258A
 ; Patent No. 6559124
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Brien, John S.
 ; APPLICANT: Kishimoto, Yasuo
 ; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
 ; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED
 ; FROM THEREFROM
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Knobbe, Martens, Olson and Bear
 ; STREET: 620 Newport Center Blvd. 16th Floor
 ; CITY: Newport Beach
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92660
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/076,258A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/958,970
FILING DATE: 28-OCT-97
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/483,146
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: 08/100,247
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.2DVIC2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
CLONE: SAPOSIN C
US-09-076-258A-3

Query Match 15.3%; Score 80; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.6e-69;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 370
DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 60
QY 371 ILLEEVSPELVCSMLHLCSG 390
DB 61 ILLEEVSPELVCSMLHLCSG 80

Query Match 15.3%; Score 80; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.6e-69;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 370
DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 60
QY 371 ILLEEVSPELVCSMLHLCSG 390
DB 61 ILLEEVSPELVCSMLHLCSG 80

Query Match 15.3%; Score 80; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.6e-69;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 370
DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 60
QY 371 ILLEEVSPELVCSMLHLCSG 390
DB 61 ILLEEVSPELVCSMLHLCSG 80

Query Match 15.3%; Score 80; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.6e-69;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 370
DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 60
QY 371 ILLEEVSPELVCSMLHLCSG 390
DB 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 14
US-08-756-031-3
Sequence 3, Application US/08756031
Patent No. 6590074
GENERAL INFORMATION:
APPLICANT: O'BRIEN, JOHN S.
APPLICANT: KISHIMOTO, YASUO
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/756,031
FILING DATE: 26-NOV-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/100,247

FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: OBRIEN.002A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
CLONE: SAPOSIN C
US-08-756-031-3

Query Match 15.3%; Score 80; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.6e-69;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 370
DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 60
QY 371 ILLEEVSPELVCSMLHLCSG 390
DB 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 15
US-08-584-671-15
Sequence 15, Application US/08584671
Patent No. 5910568
GENERAL INFORMATION:
APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,
APPLICANT: CRAWER, PALMER
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM
TITLE OF INVENTION: TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE
TITLE OF INVENTION: TO ENHANCE OR DECREASE POTENTIAL FERTILITY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA
ADDRESSEE: STATE UNIVERSITY
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/584,671
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MONAHAN, THOMAS J
REGISTRATION NUMBER: 29835
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 814-865-6277
TELEFAX: 814-865-3591
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 80

;
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: UNKNOWN
US-08-584-671-15

Query Match 14.3%; Score 75; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.7e-64;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 134 ALNLCESLQKHIAELNHQKLESNKIPELDMTEVVVAPFMANIPLLLYPQDGPRSKPQPKD 193
DB 6 ALNLCESLQKHIAELNHQKLESNKIPELDMTEVVVAPFMANIPLLLYPQDGPRSKPQPKD 65
QY 194 NGDVCQDCIQWVTDI 208
DB 66 NGDVCQDCIQWVTDI 80

Search completed: May 5, 2004, 13:35:19
Job time : 16.595 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:34:39 ; Search time 53.8081 Seconds
(without alignments)
2751.537 Million cell updates/sec

Title: US-09-743-684A-1

Perfect score: 524

Sequence: 1 MYALFLASLLGALAGPVL.....NTETAACNAVEKCRHVWN 524

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A Geneseq 29Jan04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	524	100.0	524	2 AAR70783	Aar70783 Prosaposin
2	524	100.0	524	2 AAW85652	Aaw85652 Human pro
3	524	100.0	524	3 AAY58716	Aay58716 Human pro
4	524	100.0	524	6 ABY79099	Abu79099 Lip-TAA b
5	524	100.0	524	6 ABU05200	Abu05200 Human exp
6	524	100.0	524	6 ABU05207	Abu05207 Human exp
7	524	100.0	524	6 ABU05203	Abu05203 Human exp
8	524	100.0	524	6 ABU07340	Abu07340 Human exp
9	524	100.0	524	6 ABU05216	Abu05216 Human exp
10	524	100.0	524	6 ABU05202	Abu05202 Human exp
11	524	100.0	524	6 ABU05214	Abu05214 Human exp
12	524	100.0	524	6 ABU05215	Abu05215 Human exp
13	524	100.0	524	6 ABU05199	Abu05199 Human exp
14	524	100.0	524	6 ABU05212	Abu05212 Human exp
15	524	100.0	524	6 ABU05213	Abu05213 Human exp
16	524	100.0	524	6 ABU05205	Abu05205 Human exp
17	423	80.7	524	6 ABU05208	Abu05208 Human exp
18	403	76.9	523	4 AAB31916	Aab31916 Amino aci
19	403	76.9	523	6 ABU05211	Abu05211 Human exp
20	284	54.2	527	4 AAB31915	Aab31915 Amino aci
21	284	54.2	527	5 ABP68602	Abp68602 Human pan
22	284	54.2	527	6 ABU79100	Abu79100 Lip-TAA b
23	284	54.2	527	6 ABU05204	Abu05204 Human exp
24	284	54.2	527	6 ABU05210	Abu05210 Human exp
25	265	50.6	385	6 ABR41750	Abu41750 Human DIT

ALIGNMENTS

RESULT 1

AAR70783
ID AAR70783 standard; protein; 524 AA.

XX AC AAR70783;

XX XX

DT 25-MAR-2003 DT (revised)

DT 30-AUG-1995 (first entry)

XX XX

DE Prosaposin.

XX XX

KW Saposin-C; neuron; myelination; nervous system; neuroblastoma;

KW neurotrophic peptide; multiple sclerosis; leukoencephalitis;

KW adrenal leukodystrophy; prosaposin.

XX XX

OS Homo sapiens.

XX XX

PN WO9503821-A1.

XX XX

PD 09-FEB-1995.

XX XX

PF 28-JUL-1994; 94WO-US008453.

XX XX

PR 30-JUL-1993; 93US-00100247.

PR 21-APR-1994; 94US-00232513.

XX XX

PA (OBRI/) O'BRIEN J S.

XX XX

PI O'brien JS, Kishimoto Y;

XX XX

DR WPI; 1995-082029/11.

DR N-PSDB; AAQ85355.

XX XX

PT Stimulating neural cell out-growth and myelination - with pro-saposin, saposin C or new neurotrophic peptide(s) from cytokine(s), for treating nervous system diseases.

XX XX

PS Disclosure; Page 30-32; 50pp; English.

XX XX

CC The peptide given in AAR70773, corresponding to amino acids 8-29 of human saposin-C (AAR70784), promotes neurite outgrowth in vitro. A consensus sequence was determined by comparing the peptide with hematopoietic and neurotrophic cytokines, and neurotrophic peptides (AAR70774-82) were identified in the AB loop of human ciliary neurotrophic factor. interleukins-6, -2, -3 and -gamma, erythropoietin and leukocyte inhibitory factor, and in helix C of human interleukin-1-beta and oncostatin-M. Prosaposin (AAR70783) and saposin-C also promoted nerve cell myelination ex vivo. (Updated on 25-MAR-2003 to correct PN field.)

Abu05209 Human exp
Abu05206 Human exp
Aab39442 Human GEN
Abu05201 Human exp
Abg70166 Human pre
Abu70504 Human adi
Abu70799 Human adi
Abu70422 Human adi
Aar70784 Saposin-C
Aaw85653 Human sap
Aau05697 Human Sap
Abu62252 Spingoli
Aau05698 Human glu
Aab31912 Amino aci
Aab31929 Amino aci
Aaw18584 Universal
Aay33258 Human pro
Abu62249 Spingoli
Abu62250 Spingoli
Aab67289 Human sap

CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to
 CC correct PI field.)
 XX
 SQ Sequence 524 AA;

Query Match 100.0%; Score 524; DB 2; Length 524;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTWNKPTVKS 60
 Db 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTWNKPTVKS 60
 QY 61 LPCDICXDVVTAAGDMLKDNATEEEILVYLEKTCMDLKPKNMSASCKEIVDSYLPVILDI 120
 Db 61 LPCDICXDVVTAAGDMLKDNATEEEILVYLEKTCMDLKPKNMSASCKEIVDSYLPVILDI 120
 QY 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLY 180
 Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLY 180
 QY 181 PDGGRSKPQPKDNGDVQCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGWADI 240
 Db 181 PDGGRSKPQPKDNGDVQCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGWADI 240
 QY 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPQTLVPAKVASKNVIPELVE 300
 Db 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPQTLVPAKVASKNVIPELVE 300
 QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFQKCKLPSLSEECQEV 360
 Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFQKCKLPSLSEECQEV 360
 QY 361 VDTYSSILSILLEVPSELVCSMLHLCSTRLPALTVHTQPKDGGFCVCKLVGYLD 420
 Db 361 VDTYSSILSILLEVPSELVCSMLHLCSTRLPALTVHTQPKDGGFCVCKLVGYLD 420
 QY 421 RNLEKNSFKOILAALEKGCFLDPYQKQCDQFVAEYEPVLBIELVMDPSFVCLKIG 480
 Db 421 RNLEKNSFKOILAALEKGCFLDPYQKQCDQFVAEYEPVLBIELVMDPSFVCLKIG 480
 QY 481 ACSAHKPELLGTEKICWGPSWQNTETAACNAVEHCKRHVN 524
 Db 481 ACSAHKPELLGTEKICWGPSWQNTETAACNAVEHCKRHVN 524

RESULT 2
 AAW85652
 ID AAW85652 standard; protein; 524 AA.
 XX
 AC AAW85652;
 XX
 DT 19-JUL-1999 (first entry)
 XX
 DE Human prosaposin N-terminal peptide.
 XX

Prosaposin; saposin; prosaptides; prosaposin receptor agonists; PRA;
 peripheral nervous system; central nervous system; PNS; CNS; Akt; Bcl-2;
 therapy; treatment; apoptosis; caspase; tumour necrosis factor; TNF;
 cytokine; interferon gamma; IFN; inflammation; rheumatoid arthritis;
 Crohn's disease; irritable bowel syndrome; asthma; cardiac infarction;
 congestive heart failure; multiple sclerosis;
 acute disseminated inflammatory leukoencephalitis;
 progressive multifocal leukoencephalitis;
 Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;
 ischemic heart disease; Guillain-Barre disease; alopecia; AIDS dementia;
 cerebral malaria; HTLV; neuropathy;
 inflammatory neurodegenerative disease; toxin-induced liver disease.
 Homo sapiens.
 OS
 XX
 PN WO9912559-A1.
 XX

PD 18-MAR-1999.
 XX
 PF 09-SEP-1998; 98WO-US019216.
 XX
 PR 09-SEP-1997; 97US-0058352P.
 PR 04-JUN-1998; 98US-0088129P.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI O'Brien JS;
 XX
 DR WPI; 1999-229139/19.
 DR N-PSDB; AAX06488.
 XX
 FT Use of prosaposin receptor agonist.
 XX
 PS Claim 7; Fig 2; 90pp; English.
 XX
 CC Prosaposin is a 70kDa glycoprotein which is proteolytically processed to
 CC generate saposins A, B, C and D, all of which are similar to each other
 CC and have a similar placement of six cysteines, a glycosylation site and
 CC conserved proline residues. Prosaposin, saposin C and prosaposin derived
 CC peptides (prosaptides), have therapeutic applications in promoting
 CC recovery after toxic, traumatic, myocardial ischaemic, degenerative and
 CC inherited lesions to the peripheral and central nervous system.
 CC Prosaposin receptor agonists (PRAs) inhibit proinflammatory cytokine-
 CC induced apoptosis by activation of the Ser/Thr protein kinase Akt. Akt
 CC dissociates complexes of Bcl-2 family members, such as BAD-Bcl-2,
 CC releasing Bcl-2 and its family members which inhibit caspases, thereby
 CC inhibiting apoptosis. An additional mechanism whereby PRAs inhibit
 CC apoptosis is by blocking activation of JNK, a proapoptotic signaling
 CC component. Within several minutes after binding to the receptor, PRAs
 CC block JNK activation induced by tumor necrosis factor-alpha (TNF alpha).
 CC The activation of JNK by TNF alpha is another well known mechanism for
 CC TNF alpha-induced, as well as other proinflammatory cytokine-induced
 CC apoptosis. The method can be used for inhibiting apoptosis which is
 CC caspase-mediated or induced by a proinflammatory cytokine, for example
 CC TNF alpha or interferon-gamma. It can be used for inhibiting apoptosis
 CC associated with a disorder such as e.g. rheumatoid arthritis, Crohn's
 CC disease, irritable bowel syndrome, asthma, cardiac infarction, congestive
 CC heart failure, multiple sclerosis, acute disseminated inflammatory
 CC leukoencephalitis, progressive multifocal leukoencephalitis, Alzheimer's
 CC disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's
 CC disease, ischemic heart disease, Guillain-Barre disease, traumatic brain
 CC injury, traumatic spinal cord injury, alopecia, AIDS dementia, cerebral
 CC malaria, HTLV, neuropathy, inflammatory neurodegenerative disease, and
 CC toxin-induced liver disease. This 524 N-terminal peptide of prosaposin
 CC also acts as a prosaposin receptor agonist
 XX
 SQ Sequence 524 AA;

Query Match 100.0%; Score 524; DB 2; Length 524;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTWNKPTVKS 60
 Db 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTWNKPTVKS 60
 QY 61 LPCDICXDVVTAAGDMLKDNATEEEILVYLEKTCMDLKPKNMSASCKEIVDSYLPVILDI 120
 Db 61 LPCDICXDVVTAAGDMLKDNATEEEILVYLEKTCMDLKPKNMSASCKEIVDSYLPVILDI 120
 QY 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLY 180
 Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLY 180
 QY 181 PDGGRSKPQPKDNGDVQCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGWADI 240
 Db 181 PDGGRSKPQPKDNGDVQCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGWADI 240
 QY 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPQTLVPAKVASKNVIPELVE 300
 Db 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPQTLVPAKVASKNVIPELVE 300

Db 241 CKNYISQYSEIAIQMMHMQPKICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300
Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEBCEQEV 360
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEBCEQEV 360
Qy 361 VDTYGSILSILLEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLD 420
Db 361 VDTYGSILSILLEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLD 420
Qy 421 RNLEKNTKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNTKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524

RESULT 3
AAV58716
ID AAV58716 standard; protein; 524 AA.
XX
AC AAV58716;
XX
DT 25-APR-2000 (first entry)
XX
DE Human prosaposin.
XX
KW Prosaposin; saponin B; antiangiogenic; angiogenesis inhibitor;
KW antitumour; antiproliferative; antimigratory; Kaposi's sarcoma; tumour;
KW human; therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 195..:275
FT /note= "mature saposin B"
FT Peptide 195..:205
FT /note= "specifically claimed antiangiogenic peptide of
FT Claim 23"
FT Peptide 196..:200
FT /note= "specifically claimed antiangiogenic peptide of
FT Claim 4"
XX
PN WO200002902-A1.
XX
PD 20-JAN-2000.
XX
XX 12-JUL-1999; 99WO-US015772.
XX
XX 13-JUL-1998; 98US-0092647P.
XX
PA (GILL/) GILL P S.
XX
XX Gill PS;
XX
XX WPI; 2000-171128/15.
XX
XX Saposin B derived peptides, useful as inhibitors of angiogenesis and
XX tumor growth.
XX
XX Disclosure; Page 18; 78pp; English.
XX
CC The present sequence is that of human prosaposin, a precursor of saposin
CC B. The invention is based on the discovery that saposin B, previously
CC known to be involved in the hydrolysis of sphingolipids, has potent
CC antiangiogenic and antitumour activity, and also has antiproliferative
CC and antimigratory activity against endothelial cells. This activity is
CC conserved in cryptic polypeptides as small as 5 amino acids (see AAV58684
CC -715), which can be synthetically prepared and used in vitro or in vivo
CC for the treatment of undesired angiogenesis and tumor growth, especially
CC Kaposi's sarcoma (claimed). The polypeptides can also be used in

CC conjunction with cytotoxic moieties to selectively kill certain cell
CC types, e.g. for treatment of cancer, angiofibroma, neovascular glaucoma,
CC arteriovenous malformation, nonunion fracture, arthritis and other
CC connective tissue disorders, Osler-Weber syndrome, atherosclerotic
CC plaque, psoriasis, corneal graft neovascularization, pyogenic granuloma,
CC retrolental fibroplasia, diabetic retinopathy, scleroderma, haemangioma,
CC trachoma, vascular adhesions and hypertrophic scars
XX
SQ Sequence 524 AA;
Query Match 100.0%; Score 524; DB 3; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYALFLASLLGAALAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60
Db 1 MYALFLASLLGAALAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60
Qy 61 LPCDICKOVVTAAGDMLKDNATEEELVYLEKTCWMLPKPNMSASCKEIVDSYLPVLDI 120
Db 61 LPCDICKOVVTAAGDMLKDNATEEELVYLEKTCWMLPKPNMSASCKEIVDSYLPVLDI 120
Qy 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLILY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLILY 180
Qy 181 PODGPRSKPQKONGDVQCDCIQWVTDIQTAVRTNSTFFQALVEHVKEECDRLGPGWADI 240
Db 181 PODGPRSKPQKONGDVQCDCIQWVTDIQTAVRTNSTFFQALVEHVKEECDRLGPGWADI 240
Qy 241 CKNYISQYSEIAIQMMHMQPKICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300
Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEBCEQEV 360
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEBCEQEV 360
Qy 361 VDTYGSILSILLEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLD 420
Db 361 VDTYGSILSILLEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLD 420
Qy 421 RNLEKNTKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNTKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524

RESULT 4
ABU79099
ID ABU79099 standard; protein; 524 AA.
XX
AC ABU79099;
XX
DT 18-JUN-2003 (first entry)
XX
DE Lip-TAA binding protein, Prosaposin.
XX
KW Superantigen; SAG; staphylococcal enterotoxin; tumour; cancer; apoptosis;
KW gene therapy; mammalian cell receptor; cytostatic;
KW tumour associated lipid; anergy; T cell; antigen presenting cell; APC;
KW tumoricidal immunocyte; antitumour.
XX
OS Unidentified.
XX
PN US2002177551-A1.
XX
PD 28-NOV-2002.
XX
XX 30-MAY-2001; 2001US-00870759.
PF

31-MAY-2000; 2000US-0208128P.
(TERM/) TERMAN D S.
Terman DS;
WPI; 2003-361759/34.

A mammalian cell receptor, useful in the treatment of cancer by binding to tumor associated lipids where the binding induces anergy or apoptosis in T cells and antigen presenting cells.

Disclosure; Page; 167pp; English.

The invention relates to a mammalian cell receptor, useful in the treatment of cancer, which binds to tumor associated lipids and induces anergy or apoptosis in the T cells and antigen presenting cells (APCs). Also included are a mammalian cell useful in the treatment of cancer where the receptor which binds tumor associated lipids and induces cellular inactivation or death is deleted or functionally deactivated, producing (M1) a tumoricidal immunocyte population in vivo in a mammal (by allowing tumor associated lipids to contact immunocytes in which receptors for immunosuppressive fatty acids, ceramides, glycolipids, sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides, sialylated glycans, lipopeptides and proteoglycolipids are inactivated or deleted), a construct useful in the treatment of cancer comprising a superantigen (SAG) nucleotide inserted into a virus, a mammalian T cell useful in the treatment of cancer (where an adaptor protein which inhibits T cell activation by tumor associated antigens is deleted or functionally deactivated), a composition useful in the treatment of cancer (comprising a lipid raft conjugated to a superantigen), producing (M2) a tumoricidal immunocyte population ex vivo in a mammal (by allowing tumor associated lipids to contact immunocytes, in which receptors for the lipids are inactivated or deleted to produce a tumoricidal immunocyte population, and administering the tumoricidally activated lipids to contact T cells, in which adaptor proteins, which inhibit T cell activation by tumor associated antigens, are deleted or functionally deactivated to produce a tumoricidal population of T cells, and administering the tumoricidally activated T cells to the host, or allowing a superantigen-lipid raft to contact T cells ex vivo, and administering the tumoricidally activated T cells to the host), treating (M5) cancer in a mammal (by administering a lipid binding molecule which binds immunosuppressive tumor associated lipids in vivo), producing (M6) a tumoricidal T cell population in vivo in a mammal (by allowing a tumor associated antigen to contact immunocytes in which adaptor proteins which inhibit T cell activation by tumor associated antigens are deleted or functionally deactivated) and producing (M7) a tumoricidal T cell population ex vivo in a mammal comprising allowing a superantigen-lipid raft conjugate to contact immunocytes and receptors, methods and compositions are useful for treating cancers and tumours. Bacterial superantigens are co-administered or administered as fusion constructs with anti-tumour proteins or motifs. The present sequence represents a tumour antigen or a motif identifying a tumour antigen, which can be functionally deactivated in the method of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format from the US patent office website at "seqdata.uspto.gov/sequence.html?DocID=20020177551"

Sequence 524 AA;

Query Match 100.0%; Score 524; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MYALFLLASLLGAALAGPVLGKCEKTRGSVAVQNVKTSADCGAVKHCLQTVWVKPTVKS 60

Db 1 MYALFLLASLLGAALAGPVLGKCEKTRGSVAVQNVKTSADCGAVKHCLQTVWVKPTVKS 60
QY 61 LPCDICKDVVTAAGDMKONATEEEILVYLEKTCMDLPRNNSASCKEIVDSYLPVLDI 120
Db 61 LPCDICKDVVTAAGDMKONATEEEILVYLEKTCMDLPRNNSASCKEIVDSYLPVLDI 120
QY 121 IKGEMSRPGVCSALNLCESLQKHAEHLNKHQKLESNKIPELDMEVWAPPFMANIPALLY 180
Db 121 IKGEMSRPGVCSALNLCESLQKHAEHLNKHQKLESNKIPELDMEVWAPPFMANIPALLY 180
QY 181 PDGPRSKPQKNDGVCDCIQMVTDTQTAVRTNSTFTVQALVEHVKEECDRGLPGMADI 240
Db 181 PDGPRSKPQKNDGVCDCIQMVTDTQTAVRTNSTFTVQALVEHVKEECDRGLPGMADI 240
QY 241 CKNYISQYSEIATQMMHMQPKEICALVGFCDEVKEMPMQTLVPAKVASKNIPALELVE 300
Db 241 CKNYISQYSEIATQMMHMQPKEICALVGFCDEVKEMPMQTLVPAKVASKNIPALELVE 300
QY 301 PIKKEHVPKASDVYCEVCFVKEVTKLIDNNKTEKILDAFDKMCCKLPKLSSEECQEV 360
Db 301 PIKKEHVPKASDVYCEVCFVKEVTKLIDNNKTEKILDAFDKMCCKLPKLSSEECQEV 360
QY 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGLPALTVHVTQPKDGGFCEVCKLVGYLD 420
Db 361 VDTYSSILSILLEEVSPELVCSMLHLCSTGLPALTVHVTQPKDGGFCEVCKLVGYLD 420
QY 421 RNLEKNSTKQETLAALKEGCSFLPDYQKQDQFVAEVEPVLIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTKQETLAALKEGCSFLPDYQKQDQFVAEVEPVLIEILVEVMDPSFVCLKIG 480
QY 481 ACPSAHKPLLTGTEKINGSPSYWCNTETAACNAVEHCKRHWN 524
Db 481 ACPSAHKPLLTGTEKINGSPSYWCNTETAACNAVEHCKRHWN 524

RESULT 5
ABU05200
ID ABU05200 standard; protein; 524 AA.
XX ABU05200;
AC ABU05200;
DT 29-JAN-2003 (first entry)
XX Human expressed protein tag (EPT) #1866.
XX
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX Homo sapiens.
XX
XX WO200278524-A2.
XX 10-OCT-2002.
XX
XX 28-MAR-2002; 2002WO-US009671.
XX
XX 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCO INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
XX

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX

PS Example 2; SEQ ID NO 1866; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 524 AA;

Query Match 100.0%; Score 524; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALLLASLILGALAGPVLGLKCTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60
Db 1 MYALLLASLILGALAGPVLGLKCTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60
Qy 61 LPDCICKDVVTAAGDMLKONATEEIIIVYLEKTCMDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPDCICKDVVTAAGDMLKONATEEIIIVYLEKTCMDWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGMSRPGVCSALNLCESIQKHLAEINHOQLESNKIPELDMTEVVAPPMANIPILLY 180
Db 121 IKGMSRPGVCSALNLCESIQKHLAEINHOQLESNKIPELDMTEVVAPPMANIPILLY 180
Qy 181 PQDGRSKPQPKDNGDVQCQIQMTDTQAVRTNSTFVQALVEHVEECDRLGPGNADI 240
Db 181 PQDGRSKPQPKDNGDVQCQIQMTDTQAVRTNSTFVQALVEHVEECDRLGPGNADI 240
Qy 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCDVEKEMPMQTLVPAKVASKNVIPALBELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCDVEKEMPMQTLVPAKVASKNVIPALBELVE 300
Qy 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKELTDADFQKWCCKLPKSLSEECQEV 360
Db 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKELTDADFQKWCCKLPKSLSEECQEV 360
Qy 361 VDTYSSSILSLLEEVSPELVCSMLHLCSTRLPALTVHTQPKDGGFCEVCKKLVGYLD 420
Db 361 VDTYSSSILSLLEEVSPELVCSMLHLCSTRLPALTVHTQPKDGGFCEVCKKLVGYLD 420
Qy 421 RNLEKNTKQIBILAALBKGSFLPDPKQKCDQFVAEYEPVLIILVEVMDPSFVCLKIG 480
Db 421 RNLEKNTKQIBILAALBKGSFLPDPKQKCDQFVAEYEPVLIILVEVMDPSFVCLKIG 480
Qy 481 ACPSAHPLLGTEKCIWGSPVWCQNTETAACNAVEHCKRHVN 524
Db 481 ACPSAHPLLGTEKCIWGSPVWCQNTETAACNAVEHCKRHVN 524

RESULT 6

ABU05207

ID ABU05207 standard; protein; 524 AA.

XX

AC ABU05207;

XX

29-JAN-2003 (first entry)

Human expressed protein tag (EPT) #1873.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

OS Homo sapiens.

XX WO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

XX 21-MAY-2001; 2001US-0292544P.

XX 08-AUG-2001; 2001US-0310801P.

XX 01-OCT-2001; 2001US-0326370P.

XX 04-DEC-2001; 2001US-0336780P.

XX 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCOS INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.

XX Example 2; SEQ ID NO 1873; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 524 AA;

Query Match 100.0%; Score 524; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALLLASLILGALAGPVLGLKCTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60

Db 1 MYALLLASLILGALAGPVLGLKCTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60

Qy 61 LPDCICKDVVTAAGDMLKONATEEIIIVYLEKTCMDWLPKPNMSASCKEIVDSYLPVILDI 120

Db 61 LPDCICKDVVTAAGDMLKONATEEIIIVYLEKTCMDWLPKPNMSASCKEIVDSYLPVILDI 120

Qy 121 IKGMSRPGVCSALNLCESIQKHLAEINHOQLESNKIPELDMTEVVAPPMANIPILLY 180

Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWAPFMANIPLLLY 180
Qy 181 PDGPRSKPQKNGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPMADI 240
Db 181 PDGPRSKPQKNGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPMADI 240
Qy 241 CKNYISQYSIAIOMMMHQPKEICALVGFCDVEKMPQTLVPAKVASKNVIPALELVE 300
Db 241 CKNYISQYSIAIOMMMHQPKEICALVGFCDVEKMPQTLVPAKVASKNVIPALELVE 300
Qy 301 PIKKEVPAKSDVYCEVCFELVKEVTKLIDNNKTEKEILDADFKMCKSLPKSLSECOEV 360
Db 301 PIKKEVPAKSDVYCEVCFELVKEVTKLIDNNKTEKEILDADFKMCKSLPKSLSECOEV 360
Qy 361 VDTYGSLSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKLVGYLD 420
Db 361 VDTYGSLSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKLVGYLD 420
Qy 421 RNLEKNSTKQEIILAALEKGCSELPDPYQKQCDQFVAEYEPVLIIEILVEWMDPSFVCLKIG 480
Db 421 RNLEKNSTKQEIILAALEKGCSELPDPYQKQCDQFVAEYEPVLIIEILVEWMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524

RESULT 7

ABU05203
ID ABU05203 standard; protein; 524 AA.

AC ABU05203;

DT 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #1869.

DE Translational profiling; expressed protein tag; EPT; kinase; phosphatase;

XX protease; protease inhibitor; transporter; cytoskeletal protein;

KW receptor; transcription factor; cancer; MHC;

KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;

KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

OS WO200278524-A2.

PN 10-OCT-2002.

PD 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

PR 21-MAY-2001; 2001US-0292544P.

PR 08-AUG-2001; 2001US-0310801P.

PR 01-OCT-2001; 2001US-0326370P.

PR 04-DEC-2001; 2001US-0336780P.

PR 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCO INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

PI WPI; 2003-040607/03.

DR New polypeptides (e.g. kinases, phosphatases, proteases, transporters,

PT cytoskeletal proteins, receptors or transcription factors), useful for

PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or

PT leukemia.

XX Example 2; SEQ ID NO 1869; 134pp; English.

PS The invention describes a purified polypeptide, which comprises a

XX fragment of a kinase, phosphatase, protease, protease inhibitor,

CC CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 524 AA;

Qy Query Match 100.0%; Score 524; DB 6; Length 524;

Db Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALFLLASLLGAALAGPVGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60

Db 1 MYALFLLASLLGAALAGPVGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60

Qy 61 LPCDICKOVVTAAGDMLKDNATEEBILVYLEKTCOWLPKPNMSASCKEIVDSYLPVLDI 120

Db 61 LPCDICKOVVTAAGDMLKDNATEEBILVYLEKTCOWLPKPNMSASCKEIVDSYLPVLDI 120

Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWAPFMANIPLLLY 180

Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWAPFMANIPLLLY 180

Qy 181 PDGPRSKPQKNGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPMADI 240

Db 181 PDGPRSKPQKNGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPMADI 240

Qy 241 CKNYISQYSIAIOMMMHQPKEICALVGFCDVEKMPQTLVPAKVASKNVIPALELVE 300

Db 241 CKNYISQYSIAIOMMMHQPKEICALVGFCDVEKMPQTLVPAKVASKNVIPALELVE 300

Qy 301 PIKKEVPAKSDVYCEVCFELVKEVTKLIDNNKTEKEILDADFKMCKSLPKSLSECOEV 360

Db 301 PIKKEVPAKSDVYCEVCFELVKEVTKLIDNNKTEKEILDADFKMCKSLPKSLSECOEV 360

Qy 361 VDTYGSLSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKLVGYLD 420

Db 361 VDTYGSLSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKLVGYLD 420

Qy 421 RNLEKNSTKQEIILAALEKGCSELPDPYQKQCDQFVAEYEPVLIIEILVEWMDPSFVCLKIG 480

Db 421 RNLEKNSTKQEIILAALEKGCSELPDPYQKQCDQFVAEYEPVLIIEILVEWMDPSFVCLKIG 480

Qy 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524

Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524

XX RESULT 8

ABU07340

ID ABU07340 standard; protein; 524 AA.

XX AC ABU07340;

XX DT 29-JAN-2003 (first entry)

XX DE Human expressed protein tag (EPT) #2041.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;

KW protease; protease inhibitor; transporter; cytoskeletal protein;

KW receptor; transcription factor; cancer; MHC;

KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX


```

KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
OS Homo sapiens.
PN WO200278524-A2.
XX 10-OCT-2002.
XX 28-MAR-2002; 2002WO-US009671.
XX 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX (ZYCO-) ZYCOS INC.
XX Chicx RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukaemia.
XX Example 2; SEQ ID NO 2041; 134pp; English.
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX polynucleotide are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPD at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 524 AA;
XX
Query Match 100.0%; Score 524; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MYALFLASLLGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTWNKPTVKS 60
Db 1 MYALFLASLLGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTWNKPTVKS 60
QY 61 LPDCICKDVVTAGDMLKDNATESEIILVLEKTCMDLPKPNMSASCKEIVDSYLPVLDI 120
Db 61 LPDCICKDVVTAGDMLKDNATESEIILVLEKTCMDLPKPNMSASCKEIVDSYLPVLDI 120
QY 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPILLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPILLY 180
QY 181 PQDGPRSKPQKNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKKECDRLGPGNADI 240
Db 181 PQDGPRSKPQKNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKKECDRLGPGNADI 240
QY 241 CKNIYSOYSEIATQMMHMQPKEICALVGCDEVKEMPMQTLVPKAVSKKNVIPALELVE 300
Db 241 CKNIYSOYSEIATQMMHMQPKEICALVGCDEVKEMPMQTLVPKAVSKKNVIPALELVE 300
QY 301 PIKKHEVPAPKSDVYCEVCFVKEVTKLIDNNKTEKEILDAPDKMCSKLPKSLSEECQEV 360
Db 301 PIKKHEVPAPKSDVYCEVCFVKEVTKLIDNNKTEKEILDAPDKMCSKLPKSLSEECQEV 360
QY 361 VPTYGSSIIISILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCFVCKKLVGYLD 420
Db 361 VPTYGSSIIISILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCFVCKKLVGYLD 420
QY 421 RNLEKNSTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
QY 481 ACPSAHKPLGLTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524
Db 481 ACPSAHKPLGLTEKCIWGPSYWCNTETAACQNAVEHCKRHVWN 524
RESULT 9
ABU05216
ID ABU05216 standard; protein; 524 AA.
XX
AC ABU05216;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1882.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
DR New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukaemia.
XX Example 2; SEQ ID NO 1882; 134pp; English.
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX polynucleotide are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPD at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 524 AA;
XX
Query Match 100.0%; Score 524; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MYALFLASLLGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTWNKPTVKS 60
Db 1 MYALFLASLLGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTWNKPTVKS 60
QY 61 LPDCICKDVVTAGDMLKDNATESEIILVLEKTCMDLPKPNMSASCKEIVDSYLPVLDI 120
Db 61 LPDCICKDVVTAGDMLKDNATESEIILVLEKTCMDLPKPNMSASCKEIVDSYLPVLDI 120
QY 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPILLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPILLY 180
QY 181 PQDGPRSKPQKNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKKECDRLGPGNADI 240
Db 181 PQDGPRSKPQKNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKKECDRLGPGNADI 240
QY 241 CKNIYSOYSEIATQMMHMQPKEICALVGCDEVKEMPMQTLVPKAVSKKNVIPALELVE 300
Db 241 CKNIYSOYSEIATQMMHMQPKEICALVGCDEVKEMPMQTLVPKAVSKKNVIPALELVE 300

```

CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 524 AA;

Query Match 100.0%; Score 524; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLASLGAALAGPVGLGKCTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
DB 1 MYALFLASLGAALAGPVGLGKCTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
QY 61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTDWLPKPNMSASCKEIVDSYLPVLDI 120
DB 61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTDWLPKPNMSASCKEIVDSYLPVLDI 120
QY 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLY 180
DB 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLY 180
QY 181 PQDGRSKPQPKNGDVCQDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
DB 181 PQDGRSKPQPKNGDVCQDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPKIECALVGFCDVEKEMPQTLVPAKASKNVIIPALBELVE 300
DB 241 CKNYISQYSEIAIQMMHMQPKIECALVGFCDVEKEMPQTLVPAKASKNVIIPALBELVE 300
QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEV 360
DB 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEV 360
QY 361 VDTYGSLSILLEEVSPVLCVSMHLCSGTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420
DB 361 VDTYGSLSILLEEVSPVLCVSMHLCSGTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420
QY 421 RNLEKNSTKQIFLALEKSGFLPDYQKQCDQFVAEYEPVLIILVEVMDPSFVCLIKIG 480
DB 421 RNLEKNSTKQIFLALEKSGFLPDYQKQCDQFVAEYEPVLIILVEVMDPSFVCLIKIG 480
QY 481 ACPSAHPELLTEKICWGPSYWCQNTTETAQCNVAHECHKRHVWN 524
DB 481 ACPSAHPELLTEKICWGPSYWCQNTTETAQCNVAHECHKRHVWN 524

RESULT 10
ABU05202
ID ABU05202 standard; protein; 524 AA.
XX AC ABU05202;
XX XX
XX 29-JAN-2003 (first entry)
XX Human expressed protein tag (EPT) #1868.
XX DE
XX Translation profiling; expressed protein tag; EPT; kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX OS Homo sapiens.
XX XX
XX WO200278524-A2.
XX PD 10-OCT-2002.
XX XX
XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX (ZYCO-) ZYCOS INC.
FA Chicx RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
PS Example 2; SEQ ID NO 1868; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 524 AA;

Query Match 100.0%; Score 524; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLASLGAALAGPVGLGKCTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
DB 1 MYALFLASLGAALAGPVGLGKCTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
QY 61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTDWLPKPNMSASCKEIVDSYLPVLDI 120
DB 61 LPDCICKDVVTAAGDMLKDNATEEELVYLEKTDWLPKPNMSASCKEIVDSYLPVLDI 120
QY 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLY 180
DB 121 IKGMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLY 180
QY 181 PQDGRSKPQPKNGDVCQDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
DB 181 PQDGRSKPQPKNGDVCQDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPKIECALVGFCDVEKEMPQTLVPAKASKNVIIPALBELVE 300
DB 241 CKNYISQYSEIAIQMMHMQPKIECALVGFCDVEKEMPQTLVPAKASKNVIIPALBELVE 300
QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEV 360
DB 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEV 360
QY 361 VDTYGSLSILLEEVSPVLCVSMHLCSGTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420
DB 361 VDTYGSLSILLEEVSPVLCVSMHLCSGTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420

Qy 421 RNLEKSTKQEIILAALEKGCSPDPYKQCDQFVAEYEPVLIEILVEMDPSFVCLKIG 480
Db 421 RNLEKSTKQEIILAALEKGCSPDPYKQCDQFVAEYEPVLIEILVEMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524

RESULT 11
ID ABU05214 standard; protein; 524 AA.
XX ABU05214;
XX 29-JAN-2003 (first entry)
DT Human expressed protein tag (EPT) #1880.
DE Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX Homo sapiens.
OS WO200278524-A2.
XX 10-OCT-2002.
XX 28-MAR-2002; 2002WO-US009671.
XX 28-MAR-2001; 2001US-0279495P.
XX 21-MAY-2001; 2001US-0292544P.
XX 08-AUG-2001; 2001US-0310801P.
XX 01-OCT-2001; 2001US-0326370P.
XX 04-DEC-2001; 2001US-0336780P.
XX 20-FEB-2002; 2002US-0358985P.
XX (ZYCO-) ZYCOS INC.
XX Chicx RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX Example 2; SEQ ID NO 1880; 134pp; English.
PS The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 524 AA;

Query Match 100.0%; Score 524; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYALFLLSLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVNKPVTKS 60
Db 1 MYALFLLSLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVNKPVTKS 60
Qy 61 LPCDICDQVVTAGDMLKDNATEEEILVYLEKTCMDLKPKNMSASCKEIVDSYLPVLDI 120
Db 61 LPCDICDQVVTAGDMLKDNATEEEILVYLEKTCMDLKPKNMSASCKEIVDSYLPVLDI 120
Qy 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAPFMANIPLLLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAPFMANIPLLLY 180
Qy 181 PDGPRSKPQKNDGVDVDCIOMVTDIQTAVRTNSTFVOALVEHVKEECDRLGPGWADI 240
Db 181 PDGPRSKPQKNDGVDVDCIOMVTDIQTAVRTNSTFVOALVEHVKEECDRLGPGWADI 240
Qy 241 CKNYISQYSEIAIQMMHMQPKEICALVGFCDEVKEMPMQTLVPKVASKNVIPALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKEICALVGFCDEVKEMPMQTLVPKVASKNVIPALELVE 300
Qy 301 PIKHEVPKASDVYCEVCFVLKVEVTKLIDNNKTEKILDAFKMCSKLPKSLSEECQEV 360
Db 301 PIKHEVPKASDVYCEVCFVLKVEVTKLIDNNKTEKILDAFKMCSKLPKSLSEECQEV 360
Qy 361 VDTYSSIIISILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420
Db 361 VDTYSSIIISILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420
Qy 421 RNLEKSTKQEIILAALEKGCSPDPYKQCDQFVAEYEPVLIEILVEMDPSFVCLKIG 480
Db 421 RNLEKSTKQEIILAALEKGCSPDPYKQCDQFVAEYEPVLIEILVEMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524

RESULT 12
ABU05215
ID ABU05215 standard; protein; 524 AA.
XX ABU05215;
XX AC ABU05215;
DT 29-JAN-2003 (first entry)
XX Human expressed protein tag (EPT) #1881.
DE Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX Homo sapiens.
OS WO200278524-A2.
XX 10-OCT-2002.
XX 28-MAR-2002; 2002WO-US009671.
XX 28-MAR-2001; 2001US-0279495P.
XX 21-MAY-2001; 2001US-0292544P.
XX 08-AUG-2001; 2001US-0310801P.
XX 01-OCT-2001; 2001US-0326370P.
XX 04-DEC-2001; 2001US-0336780P.
XX 20-FEB-2002; 2002US-0358985P.
XX (ZYCO-) ZYCOS INC.
PA

```
XX Chicz RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1881; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 524 AA;
XX
XX Query Match 100.0%; Score 524; DB 6; Length 524;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MYAFLLASLLGAALAGPVLGKCTRGSAVQCQNVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 1 MYAFLLASLLGAALAGPVLGKCTRGSAVQCQNVKTASDCGAVKHCLQTVWVKPTVKS 60
QY 61 LPDCICKDVVTAAGDMLKDNATEBEILLVLEKTCMDLWPKNMSASCKEIVDSYLPVILDI 120
Db 61 LPDCICKDVVTAAGDMLKDNATEBEILLVLEKTCMDLWPKNMSASCKEIVDSYLPVILDI 120
QY 121 IKGMSRPGVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVVAFFMANIPILLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVVAFFMANIPILLY 180
QY 181 PQDGRSKPQPKNGDVQCDCIQMVTDIQTAVRTNSTFVQALVHVHKEECDRILGPGMADI 240
Db 181 PQDGRSKPQPKNGDVQCDCIQMVTDIQTAVRTNSTFVQALVHVHKEECDRILGPGMADI 240
QY 241 CKNVISQYSEIATIQMMHMQPKETCALVGFCDVEKEMPQTLVPAKVASKNVIPALELVE 300
Db 241 CKNVISQYSEIATIQMMHMQPKETCALVGFCDVEKEMPQTLVPAKVASKNVIPALELVE 300
QY 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFKMCSKLPKSLSECCQV 360
Db 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFKMCSKLPKSLSECCQV 360
QY 361 VDTYSSSILSILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLD 420
Db 361 VDTYSSSILSILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLD 420
QY 421 RNLEKNSKTEIILAALEKSGSFDPDYQKQDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSKTEIILAALEKSGSFDPDYQKQDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480
QY 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVWN 524
```

```
RESULT 13
ABU05199
ID ABU05199 standard; protein; 524 AA.
XX
XX AC ABU05199;
XX
XX DT 29-JAN-2003 (first entry)
XX
XX DE Human expressed protein tag (EPT) #1865.
XX
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
XX OS Homo sapiens.
XX
XX PN WC200278524-A2.
XX
XX PD 10-OCT-2002.
XX
XX PF 28-MAR-2002; 2002WO-US009671.
XX
XX PR 28-MAR-2001; 2001US-0279495P.
XX
XX PR 21-MAY-2001; 2001US-0292544P.
XX
XX PR 08-AUG-2001; 2001US-0310801P.
XX
XX PR 01-OCT-2001; 2001US-0326370P.
XX
XX PR 04-DEC-2001; 2001US-0336780P.
XX
XX PR 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOS INC.
XX
XX Chicz RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
XX DR New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX PS Example 2; SEQ ID NO 1865; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 524 AA;
XX
XX Query Match 100.0%; Score 524; DB 6; Length 524;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MYAFLLASLLGAALAGPVLGKCTRGSAVQCQNVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 1 MYAFLLASLLGAALAGPVLGKCTRGSAVQCQNVKTASDCGAVKHCLQTVWVKPTVKS 60
QY 61 LPDCICKDVVTAAGDMLKDNATEBEILLVLEKTCMDLWPKNMSASCKEIVDSYLPVILDI 120
Db 61 LPDCICKDVVTAAGDMLKDNATEBEILLVLEKTCMDLWPKNMSASCKEIVDSYLPVILDI 120
```

Db 61 LPCDICKVVTAAAGDMLKDNATEEEILVYLEKTCQDMLPKFNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGEMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPMANIPLLY 180
Db 121 IKGEMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPMANIPLLY 180
Qy 181 PQDGRSKPQKNDGVCQDCIQMVTDIQTAVRTNSTFFQALVEHVKEECDRGLPGWADI 240
Db 181 PQDGRSKPQKNDGVCQDCIQMVTDIQTAVRTNSTFFQALVEHVKEECDRGLPGWADI 240
Qy 241 CKNYISOYSIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300
Db 241 CKNYISOYSIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300
Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSKLPKSLSEECQEV 360
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSKLPKSLSEECQEV 360
Qy 361 VDTYGSILSILLEEVSPELVCSMLHLCSTGTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420
Db 361 VDTYGSILSILLEEVSPELVCSMLHLCSTGTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420
Qy 421 RNLEKSTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIILVEVMDPSFVCLKIG 480
Db 421 RNLEKSTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIILVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524

RESULT 14
ABU05212
ID ABU05212 standard; protein; 524 AA.
AC ABU05212;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1878.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN W0200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCO INC.
XX
PI Chicx RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.

XX
PS
XX
CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 524 AA;

Query Match 100.0%; Score 524; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MYALFLLASLLGAALAGPVGLKCTRGSAVWCNVKTASDCGAVKHCLQTVNKNPTVKS 60
Db 1 MYALFLLASLLGAALAGPVGLKCTRGSAVWCNVKTASDCGAVKHCLQTVNKNPTVKS 60
Qy 61 LPCDICKVVTAAAGDMLKDNATEEEILVYLEKTCQDMLPKFNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKVVTAAAGDMLKDNATEEEILVYLEKTCQDMLPKFNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGEMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPMANIPLLY 180
Db 121 IKGEMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPMANIPLLY 180
Qy 181 PQDGRSKPQKNDGVCQDCIQMVTDIQTAVRTNSTFFQALVEHVKEECDRGLPGWADI 240
Db 181 PQDGRSKPQKNDGVCQDCIQMVTDIQTAVRTNSTFFQALVEHVKEECDRGLPGWADI 240
Qy 241 CKNYISOYSIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300
Db 241 CKNYISOYSIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300
Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSKLPKSLSEECQEV 360
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSKLPKSLSEECQEV 360
Qy 361 VDTYGSILSILLEEVSPELVCSMLHLCSTGTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420
Db 361 VDTYGSILSILLEEVSPELVCSMLHLCSTGTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420
Qy 421 RNLEKSTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIILVEVMDPSFVCLKIG 480
Db 421 RNLEKSTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIILVEVMDPSFVCLKIG 480
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524
Db 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524

RESULT 15
ABU05213
ID ABU05213 standard; protein; 524 AA.
XX
AC ABU05213;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1879.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.
XX Homo sapiens.
XX OS
XX WC200278524-A2.
XX 10-OCT-2002.
XX 28-MAR-2002; 2002WO-US009671.
XX 28-MAR-2001; 2001US-0279495P.
XX 21-MAY-2001; 2001US-0292544P.
XX 08-AUG-2001; 2001US-0310801P.
XX 01-OCT-2001; 2001US-0326370P.
XX 04-DEC-2001; 2001US-0336780P.
XX 20-FEB-2002; 2002US-0358985P.
XX (ZYCO-). ZYCO INC.
XX Chicz RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX Example 2; SEQ ID NO 1879; 134pp; English.
XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 524 AA;
SQ
Query Match 100.0%; Score 524; DB 6; Length 524;
Best Local Similarity 100.0%; Pred No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MYALFLASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKNPTVKS 60
DB 1 MYALFLASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKNPTVKS 60
QY 61 LPDCIDKDVVTAAGDMLKNATETEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
DB 61 LPDCIDKDVVTAAGDMLKNATETEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
QY 121 IKGEMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVFNNIPLLY 180
DB 121 IKGEMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVFNNIPLLY 180
QY 181 PQDGRSKPOPKONGDVCDICQVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240
DB 181 PQDGRSKPOPKONGDVCDICQVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240

Db 181 PQDGRSKPOPKONGDVCDICQVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240
QY 241 CKNYISOYSIAIQMMHMQPKKEICALVGFCDVEKMPQTLVPKVASKNVIPAELVE 300
Db 241 CKNYISOYSIAIQMMHMQPKKEICALVGFCDVEKMPQTLVPKVASKNVIPAELVE 300
QY 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMGSKLPKSLSEECQEV 360
Db 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMGSKLPKSLSEECQEV 360
QY 361 VDTYGSILSILLEEVSPELVCSMLHLCSTRLPALTVHVTOPKDGGFCEVCKLVGYLD 420
Db 361 VDTYGSILSILLEEVSPELVCSMLHLCSTRLPALTVHVTOPKDGGFCEVCKLVGYLD 420
QY 421 RNLEKNSTKQEILAALEKGCSEFLPDYPYKQCQDFVAEYEPVLIETILVEVMDPSFVCLKIG 480
Db 421 RNLEKNSTKQEILAALEKGCSEFLPDYPYKQCQDFVAEYEPVLIETILVEVMDPSFVCLKIG 480
QY 481 ACPSAHKPLLLGTEKCTIWGSPSYWCQNTETAAQCNAVEHCHKRVWN 524
Db 481 ACPSAHKPLLLGTEKCTIWGSPSYWCQNTETAAQCNAVEHCHKRVWN 524

Search completed: May 5, 2004, 13:45:40
Job time : 54.8081 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:34:39 ; Search time 53.1919 Seconds
(without alignments)
2751.537 Million cell updates/sec

Title: US-09-743-684a-1_COPY_7_524

Perfect score: 518

Sequence: 1 LASLLGALAGPVLGLKCT.....NTETAACNAVEHCKRHWN 518

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	518	100.0	524	2	AAR70783
2	518	100.0	524	2	AAW85652
3	518	100.0	524	3	AAV58716
4	518	100.0	524	6	ABU79099
5	518	100.0	524	6	ABU05200
6	518	100.0	524	6	ABU05207
7	518	100.0	524	6	ABU05203
8	518	100.0	524	6	ABU07340
9	518	100.0	524	6	ABU05216
10	518	100.0	524	6	ABU05202
11	518	100.0	524	6	ABU05214
12	518	100.0	524	6	ABU05215
13	518	100.0	524	6	ABU05199
14	518	100.0	524	6	ABU05212
15	518	100.0	524	6	ABU05213
16	518	100.0	524	6	ABU05205
17	417	80.5	524	6	ABU05208
18	397	76.6	523	4	AAH31916
19	397	76.6	523	6	ABU05211
20	278	53.7	527	4	AAH31915
21	278	53.7	527	5	ABP68602
22	278	53.7	527	6	ABU79100
23	278	53.7	527	6	ABU05204
24	278	53.7	527	6	ABU05210
25	265	51.2	385	6	ABR41750

26	265	51.2	526	6	ABU05209	Abu05209 Human exp
27	265	51.2	527	6	ABU05206	Abu05206 Human exp
28	254	49.0	479	6	ABR39442	AbR39442 Human GEN
29	210	40.5	210	6	ABU05201	Abu05201 Human exp
30	209	40.3	209	5	ABG70166	ABG70166 Human pre
31	153	29.5	153	6	ABU70504	Abu70504 Human adi
32	153	29.5	153	6	ABU70799	Abu70799 Human adi
33	129	24.9	129	6	ABU70422	Abu70422 Human adi
34	80	15.4	80	2	AAR70784	AAR70784 Saposin-C
35	80	15.4	80	2	AAW85653	AAW85653 Human sap
36	80	15.4	80	4	AAU05697	AAU05697 Human Sap
37	80	15.4	80	7	ABU62252	ABU62252 Springoli
38	80	15.4	592	4	AAU05698	AAU05698 Human glu
39	68	13.1	85	4	AAH31912	AAH31912 Amino aci
40	66	12.7	83	4	AAH31929	AAH31929 Amino aci
41	61	11.8	61	2	AAW18584	AAW18584 Universal
42	61	11.8	69	2	AAV33258	AAV33258 Human pro
43	40	7.7	40	7	ABU62249	ABU62249 Springoli
44	38	7.3	38	7	ABU62250	ABU62250 Springoli
45	25	4.8	25	4	AAH67289	AAH67289 Human sap

ALIGNMENTS

RESULT 1

AAR70783
ID AAR70783 standard; protein; 524 AA.

AC AAR70783;

DT 25-MAR-2003 (revised)

DT 30-AUG-1995 (first entry)

XX Prosaposin.

KW Saposin-C; neuron; myelination; nervous system; neuroblastoma;
KW neurotrophic peptide; multiple sclerosis; leukoencephalitis;
KW adrenal leukodystrophy; prosaposin.

XX Homo sapiens.

OS WO9503821-A1.

PN 09-FEB-1995.

XX 28-JUL-1994; 94WO-US008453.

XX 30-JUL-1993; 93US-00100247.

XX 21-APR-1994; 94US-00232513.

XX (OBRI/) O'BRIEN J S.

XX O'brien JS, Kishimoto Y;

XX WPI; 1995-082029/11.

XX N-PSDB; AAQ85355.

XX Stimulating neural cell out-growth and myelination - with pro:saposin,
XX saposin C or new neurotrophic peptide(s) from cytokine(s), for treating
XX nervous system diseases.

XX Disclosure; Page 30-32; 50pp; English.

XX The peptide given in AAR70773, corresponding to amino acids 8-29 of human
XX saposin-C (AAR70784), promotes neurite outgrowth in vitro. A consensus
XX sequence was determined by comparing the peptide with hematopoietic and
XX neurotrophic cytokines, and neurotrophic peptides (AAR70774-82) were
XX identified in the AB loop of human ciliary neurotrophic factor,
XX interleukin-6, -2, -3 and -gamma, erythropoietin and leukocyte
XX inhibitory factor, and in helix C of human interleukin-1-beta and
XX oncostatin-M. Prosaposin (AAR70783) and saposin-C also promoted nerve
XX cell myelination ex vivo. (Updated on 25-MAR-2003 to correct PN field.)

CC	(Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)		Sequence 524 AA;
XX	Query Match	100.0%; Score 518; DB 2; Length 524;	
XX	Best Local Similarity	100.0%; Pred. No. 0;	
XX	Matches 518; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
QY	1	LASLLGAALAGPVGLGKETRGSAYWCQNVKTASDCGAVKHCLQTVWVKPTVKSLLPCDIC 60	
DB	7	LASLLGAALAGPVGLGKETRGSAYWCQNVKTASDCGAVKHCLQTVWVKPTVKSLLPCDIC 66	
QY	61	KOVVTAAGDMLKDNATEEPILVLEKTDWLPKPNKMSASCKEIVDSYLPVLIDIIKGMS 120	
DB	67	KOVVTAAGDMLKDNATEEPILVLEKTDWLPKPNKMSASCKEIVDSYLPVLIDIIKGMS 126	
QY	121	RGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLYPQDQPR 180	
DB	127	RGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLYPQDQPR 186	
QY	181	SKPQPKDNGDVQCDCIQWVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGWADICKNYS 240	
DB	187	SKPQPKDNGDVQCDCIQWVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGWADICKNYS 246	
QY	241	QYSETAIQMMHMQPKETCALVGFCDVEKEMPOTLVPKAVSKXNVIPALELVEPIKKHE 300	
DB	247	QYSETAIQMMHMQPKETCALVGFCDVEKEMPOTLVPKAVSKXNVIPALELVEPIKKHE 306	
QY	301	VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDMCKSKLPKLSSECEQVVDYTG 360	
DB	307	VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDMCKSKLPKLSSECEQVVDYTG 366	
QY	361	SILSILLBEVSPELVCSMLHLCSTGLPALTVHVTQPKDGFCEVCKLVGYLDRLNLEKN 420	
DB	367	SILSILLBEVSPELVCSMLHLCSTGLPALTVHVTQPKDGFCEVCKLVGYLDRLNLEKN 426	
QY	421	STKQEILAALEKSGSELPDPYQKQCDQFVAEYEPVLIEILVEWMDPSFVCLKTGACPSAH 480	
DB	427	STKQEILAALEKSGSELPDPYQKQCDQFVAEYEPVLIEILVEWMDPSFVCLKTGACPSAH 486	
QY	481	KPLLGTETKCIWGPSYWCQNTETATAQCNVAHECKRHVWN 518	
DB	487	KPLLGTETKCIWGPSYWCQNTETATAQCNVAHECKRHVWN 524	
RESULT 2			
AAW85652			
ID	AAW85652	standard; protein; 524 AA.	
XX	AAW85652;		
AC			
XX			
DT	19-JUL-1999	(first entry)	
XX			
DE	Human prosaposin N-terminal peptide.		
XX			
XX	Prosaposin; saposin; prosapetides; prosaposin receptor agonists; PRA;		
KW	peripheral nervous system; central nervous system; PNS; CNS; Akt; Bcl-2;		
KW	therapy; treatment; apoptosis; caspase; tumour necrosis factor; TNF;		
KW	cytokine; interferon gamma; IFN; inflammation; rheumatoid arthritis;		
KW	Crohn's disease; irritable bowel syndrome; asthma; cardiac infarction;		
KW	congestive heart failure; multiple sclerosis;		
KW	acute disseminated inflammatory leukoencephalitis;		
KW	progressive multifocal leukoencephalitis;		
KW	Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;		
KW	ischemic heart disease; Guillain-Barre disease; alopecia; AIDS dementia;		
KW	cerebral malaria; HTLV; neuropathy;		
KW	inflammatory neurodegenerative disease; toxin-induced liver disease.		
XX			
OS	Homo sapiens.		
XX			
PN	WO912559-A1.		
XX			

Db 247 QYSEIAIOMMHQPKKEICALVGFCDCEVKEMPMQTLVPAKVASKNVIPALVELPEIKKHE 306
Qy 301 VPAKSDVYCEVFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSESCQEVVDITYGS 360
Db 307 VPAKSDVYCEVFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSESCQEVVDITYGS 366
Qy 361 SILSILLEEVSPVLVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKKLVGYLDRLNLEKN 420
Db 367 SILSILLEEVSPVLVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKKLVGYLDRLNLEKN 426
Qy 421 STKQEIILAALEKGCSEFLPDYPKQCDQFVAEYEPVLEILVEVMDPSFVCLKIGACPSAH 480
Db 427 STKQEIILAALEKGCSEFLPDYPKQCDQFVAEYEPVLEILVEVMDPSFVCLKIGACPSAH 486
Qy 481 KPLLGTCKCIWGPSYWCNTETAAQCNVAHECKRHVWN 518
Db 487 KPLLGTCKCIWGPSYWCNTETAAQCNVAHECKRHVWN 524

RESULT 3
AA58716
ID AA58716 standard; protein; 524 AA.
AC AA58716;
XX
XX
DT 25-APR-2000 (first entry)
XX
DE Human prosaposin.
XX
KW Prosaposin; saponin B; antiangiogenic; angiogenesis inhibitor;
KW anticancer; antiproliferative; antimigratory; Kaposi's sarcoma; tumour;
KW human; therapy.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Protein 195..275
FT /note= "mature saposin B"
FT Peptide 195..205
FT /note= "specifically claimed antiangiogenic peptide of
FT Claim 23"
FT Peptide 196..200
FT /note= "specifically claimed antiangiogenic peptide of
FT Claim 4"
XX
XX WO200002902-A1.
XX
XX 20-JAN-2000.
XX
XX 12-JUL-1999; 99WO-US015772.
XX
XX 13-JUL-1998; 98US-0092647P.
XX
XX (GILL/) GILL P S.
XX
XX Gill PS;
XX
XX WPI; 2000-171128/15.
XX
XX Saposin B derived peptides, useful as inhibitors of angiogenesis and
XX tumor growth.
XX
XX Disclosure; Page 18; 78pp; English.
XX
XX The present sequence is that of human prosaposin, a precursor of saposin
XX B. The invention is based on the discovery that saposin B, previously
XX known to be involved in the hydrolysis of sphingolipids, has potent
XX antiangiogenic and antitumor activity, and also has antiproliferative
XX and antimigratory activity against endothelial cells. This activity is
XX conserved in cryptic polypeptides as small as 5 amino acids (see AAY58684
XX -715), which can be synthetically prepared and used in vitro or in vivo
XX for the treatment of undesired angiogenesis and tumor growth, especially
XX Kaposi's sarcoma (claimed). The polypeptides can also be used in

CC conjunction with cytotoxic moieties to selectively kill certain cell
CC types, e.g. for treatment of cancer, angiofibroma, neovascular glaucoma,
CC arteriovenous malformation, nonunion fracture, arthritis and other
CC connective tissue disorders, Osler-Weber syndrome, atherosclerotic
CC plaque, psoriasis, corneal graft neovascularization, pyogenic granuloma,
CC retrolental fibroplasia, diabetic retinopathy, scleroderma, haemangioma,
CC trachoma, vascular adhesions and hypertrophic scars
XX
SQ Sequence 524 AA;
Query Match 100.0%; Score 518; DB 3; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LASILGAALAGPVILGLKECTRGSAVWQNVKTASDCGAVKHCLQTVNMKPTVKSPLCDIC 60
Db 7 LASILGAALAGPVILGLKECTRGSAVWQNVKTASDCGAVKHCLQTVNMKPTVKSPLCDIC 66
Qy 61 KDVTAAAGDMLKDNATEEELVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIIGEMS 120
Db 67 KDVTAAAGDMLKDNATEEELVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIIGEMS 126
Qy 121 RPGEVCSALNLCESLOKHLAELNHOKLESNKIPELDMTEVAVPFMANIPLLLYPQDGP 180
Db 127 RPGEVCSALNLCESLOKHLAELNHOKLESNKIPELDMTEVAVPFMANIPLLLYPQDGP 186
Qy 181 SKPQKNDGDCVDCDCIQMVTIDIQTAVRTNSTFTVQALVEHVKECDRLGPGMADICKNYIS 240
Db 187 SKPQKNDGDCVDCDCIQMVTIDIQTAVRTNSTFTVQALVEHVKECDRLGPGMADICKNYIS 246
Qy 241 QYSEIAIOMMHQPKKEICALVGFCDCEVKEMPMQTLVPAKVASKNVIPALVELPEIKKHE 300
Db 247 QYSEIAIOMMHQPKKEICALVGFCDCEVKEMPMQTLVPAKVASKNVIPALVELPEIKKHE 306
Qy 301 VPAKSDVYCEVFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSESCQEVVDITYGS 360
Db 307 VPAKSDVYCEVFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSESCQEVVDITYGS 366
Qy 361 SILSILLEEVSPVLVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKKLVGYLDRLNLEKN 420
Db 367 SILSILLEEVSPVLVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKKLVGYLDRLNLEKN 426
Qy 421 STKQEIILAALEKGCSEFLPDYPKQCDQFVAEYEPVLEILVEVMDPSFVCLKIGACPSAH 480
Db 427 STKQEIILAALEKGCSEFLPDYPKQCDQFVAEYEPVLEILVEVMDPSFVCLKIGACPSAH 486
Qy 481 KPLLGTCKCIWGPSYWCNTETAAQCNVAHECKRHVWN 518
Db 487 KPLLGTCKCIWGPSYWCNTETAAQCNVAHECKRHVWN 524

RESULT 4
ABU79099
ID ABU79099 standard; protein; 524 AA.
XX
XX AC ABU79099;
XX
XX 18-JUN-2003 (first entry)
DT
XX
XX Lip-TMA binding protein, Prosaposin.
DE
XX
XX Superantigen; SAG; staphylococcal enterotoxin; tumour; cancer; apoptosis;
KW gene therapy; mammalian cell receptor; cytostatic;
KW tumour associated lipid; anergy; T cell; antigen presenting cell; APC;
KW tumouricidal immunocyte; antitumour.
XX
XX Unidentified.
XX
XX US2002177551-A1.
XX
XX 28-NOV-2002.
XX
XX 30-MAY-2001; 2001US-00870759.
PF

XX PR 31-MAY-2000; 2000US-0208128P.
 XX PA (TERM/) TERMAN D S.
 XX PI Terman DS;
 XX DR WPI; 2003-361759/34.
 XX PT A mammalian cell receptor, useful in the treatment of cancer by binding
 PT to tumor associated lipids where the binding induces anergy or apoptosis
 PT in T cells and antigen presenting cells.
 XX PS Disclosure; Page; 167pp; English.
 XX SS The invention relates to a mammalian cell receptor, useful in the
 CC treatment of cancer, which binds to tumor associated lipids and induces
 CC anergy or apoptosis in the T cells and antigen presenting cells (APCs).
 CC Also included are a mammalian cell useful in the treatment of cancer
 CC where the receptor which binds tumor associated lipids and induces
 CC cellular inactivation or death is deleted or functionally deactivated,
 CC producing (M1) a tumouricidal immunocyte population in vivo in a mammal
 CC (by allowing tumor associated lipids to contact immunocytes in which
 CC receptors for immunosuppressive fatty acids, ceramides, glycolipids,
 CC sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides,
 CC sialylated glycans, lipopeptides and proteoglycolipids are inactivated or
 CC deleted), a construct useful in the treatment of cancer comprising a
 CC superantigen (SAG) nucleotide inserted into a virus, a mammalian T cell
 CC useful in the treatment of cancer (where an adaptor protein which
 CC inhibits T cell activation by tumor associated antigens is deleted or
 CC functionally deactivated), a composition useful in the treatment of
 CC cancer (comprising a lipid raft conjugated to a superantigen), producing
 CC (M2) a tumouricidal immunocyte population ex vivo in a mammal (by
 CC allowing tumor associated lipids to contact immunocytes, in which
 CC receptors for the lipids are inactivated or deleted to produce a
 CC tumouricidal immunocyte population, and administering the tumouricidal
 CC activated immunocytes to the host), producing (M3) a tumouricidal APC
 CC population ex vivo in a mammal (by allowing a tumor associated lipid to
 CC contact APCs, in which receptors for the tumor associated lipid are
 CC inactivated or deleted to produce a tumouricidal activated lipids are
 CC and administering APCs to the host), producing a tumouricidal T cell
 CC population ex vivo in a mammal (by allowing a tumour associated lipid to
 CC contact T cells, in which adaptor proteins, which inhibit T cell
 CC activation by tumor associated antigens, are deleted or functionally
 CC deactivated to produce a tumouricidal population of T cells, and
 CC administering the tumouricidal activated T cells to the host, or
 CC allowing a superantigen-lipid raft to contact T cells ex vivo, and
 CC administering the tumouricidal activated T cells to the host), treating
 CC (M5) cancer in a mammal (by administering a lipid binding molecule which
 CC binds immunosuppressive tumor associated lipids in vivo), producing (M6)
 CC a tumouricidal T cell population in vivo in a mammal (by allowing a
 CC tumor associated antigen to contact immunocytes in which adaptor
 CC proteins which inhibit T cell activation by tumor associated antigens
 CC are deleted or functionally deactivated) and producing (M7) a
 CC tumouricidal T cell population ex vivo in a mammal comprising allowing a
 CC superantigen-lipid raft conjugate to contact immunocytes in vivo. The
 CC receptors, methods and compositions are useful for treating cancers and
 CC tumours. Bacterial superantigens are co-administered or administered as
 CC fusion constructs with anti-tumour proteins or motifs. The present
 CC sequence represents a tumour antigen or a motif identifying a tumour
 CC antigen, which can be functionally deactivated in the method of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format from the
 CC US patent office website at
 CC "seqdata.uspto.gov/sequence.html?DocID=20020177551"
 XX Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LASLGAALAGPVILGKCTRGSAVWQNVKTASDCGAVKHCLQTVWKNKPTVKSIPCDIC 60

Db	7	LASLGAALAGPVILGKCTRGSAVWQNVKTASDCGAVKHCLQTVWKNKPTVKSIPCDIC	66
QY	61	KDVTAAAGDMLKONATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKIGEMS	120
Db	67	KDVTAAAGDMLKONATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKIGEMS	126
QY	121	RPGEVCSALNCSLQKHLAELNHQKLESNKIPELDMDTEWVAFPMANIPLLIYPQDGP	180
Db	127	RPGEVCSALNCSLQKHLAELNHQKLESNKIPELDMDTEWVAFPMANIPLLIYPQDGP	186
QY	181	SKPQKXNDGVDCQDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPMADICKNYIS	240
Db	187	SKPQKXNDGVDCQDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPMADICKNYIS	246
QY	241	OYSEIAIQMMHMQPKEICALVGFCDCEVKEMPMQTLVPAKVASKNVIPALELVEPIKKHE	300
Db	247	OYSEIAIQMMHMQPKEICALVGFCDCEVKEMPMQTLVPAKVASKNVIPALELVEPIKKHE	306
QY	301	VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKLSSECEVVDVDTGS	360
Db	307	VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKLSSECEVVDVDTGS	366
QY	361	SILSILLESVPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLDRNLEKN	420
Db	367	SILSILLESVPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLDRNLEKN	426
QY	421	STKQETLAALKGCSTFLPDPYKQCDQFVAEVEPVLIETLVEMDPSFVCLKIGACPSAH	480
Db	427	STKQETLAALKGCSTFLPDPYKQCDQFVAEVEPVLIETLVEMDPSFVCLKIGACPSAH	486
QY	481	KPLLGTCKIMGSPSWQNTETAACNAVEHCKRHVN	518
Db	487	KPLLGTCKIMGSPSWQNTETAACNAVEHCKRHVN	524

RESULT 5
 ABU05200
 ID ABU05200 standard; protein; 524 AA.
 XX ABU05200;
 AC ABU05200;
 XX 29-JAN-2003 (first entry)
 XX Human expressed protein tag (EPT) #1866.
 DE Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 XX Homo sapiens.
 OS WO200278524-A2.
 PN 10-OCT-2002.
 PD 28-MAR-2002; 2002WO-US009671.
 PF 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX (ZYCO-) ZYCOs INC.
 PA Chicx RM, Tomlinson AJ, Urban RG;
 PI WPI; 2003-040607/03.
 DR
 XX

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.

XX Example 2; SEQ ID NO 1866; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LASLGAALAGPVGLGKCECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSLPCDIC 60
Db 7 LASLGAALAGPVGLGKCECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSLPCDIC 66
QY 61 KDVVTAAGDMLKDNATEEILVYLEKTCMDLPKNMSASCKEIVDSYLPVLDIIGKMS 120
Db 67 KDVVTAAGDMLKDNATEEILVYLEKTCMDLPKNMSASCKEIVDSYLPVLDIIGKMS 126
QY 121 RPGEVCSALNLCESIQKHLAEINHQKLESNKIPELDMTEVVAFPMANIPLLLPDQGP 180
Db 127 RPGEVCSALNLCESIQKHLAEINHQKLESNKIPELDMTEVVAFPMANIPLLLPDQGP 186
QY 181 SKPQKNDGVQDCIQMVTDIQTAVRTNSTFVQALVEHVKCECDRLGPGMADICKNYIS 240
Db 187 SKPQKNDGVQDCIQMVTDIQTAVRTNSTFVQALVEHVKCECDRLGPGMADICKNYIS 246
QY 241 QYSEIAIQMMHMOPKEICALVGFCDVEKPMQTLVPAKVASKNVIPALELVEPIKHE 300
Db 247 QYSEIAIQMMHMOPKEICALVGFCDVEKPMQTLVPAKVASKNVIPALELVEPIKHE 306
QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDITYGS 360
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDITYGS 366
QY 361 SILSILLEVSPELVCSMLHLCSGFRLPALTVHTVQPKDGGFCEVCKKLGVYLDRLNLEKN 420
Db 367 SILSILLEVSPELVCSMLHLCSGFRLPALTVHTVQPKDGGFCEVCKKLGVYLDRLNLEKN 426
QY 421 STKQELIALLKSGCSFLDPYQKCDQFVAEYEPVLILVEVMDPPSVCLIKIGACPSAH 480
Db 427 STKQELIALLKSGCSFLDPYQKCDQFVAEYEPVLILVEVMDPPSVCLIKIGACPSAH 486
QY 481 KPLLGTEKICWGPSYWCQNTETAACNAVEHCKRHVN 518
Db 487 KPLLGTEKICWGPSYWCQNTETAACNAVEHCKRHVN 524

RESULT 6
ID ABU05207
XX ABU05207 standard; protein; 524 AA.
AC ABU05207;

XX 29-JAN-2003 (first entry)
DT Human expressed protein tag (EPT) #1873.
DE
XX
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
XX WO200278524-A2.
XX 10-OCT-2002.
XX
XX 28-MAR-2002; 2002WO-US009671.
XX
XX 28-MAR-2001; 2001US-0279495P.
XX 21-MAY-2001; 2001US-0292544P.
XX 08-AUG-2001; 2001US-0310801P.
XX 01-OCT-2001; 2001US-0326370P.
XX 04-DEC-2001; 2001US-0336780P.
XX 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOS INC.

XX Chicx RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.

XX Example 2; SEQ ID NO 1873; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LASLGAALAGPVGLGKCECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSLPCDIC 60
Db 7 LASLGAALAGPVGLGKCECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSLPCDIC 66
QY 61 KDVVTAAGDMLKDNATEEILVYLEKTCMDLPKNMSASCKEIVDSYLPVLDIIGKMS 120
Db 67 KDVVTAAGDMLKDNATEEILVYLEKTCMDLPKNMSASCKEIVDSYLPVLDIIGKMS 126
QY 121 RPGEVCSALNLCESIQKHLAEINHQKLESNKIPELDMTEVVAFPMANIPLLLPDQGP 180

Db 127 RPGEVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVWAPFMANIPILLYPQDGR 186
 QY 181 SKPQPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADICKNYIS 240
 Db 187 SKPQPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADICKNYIS 246
 QY 241 QYSEIAIOMMHMPKEICALVGFCDVEKMPQTLVPAKVASKNVIIPALELVEPIKKE 300
 Db 247 QYSEIAIOMMHMPKEICALVGFCDVEKMPQTLVPAKVASKNVIIPALELVEPIKKE 306
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSECEQVWDYTG 360
 Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSECEQVWDYTG 366
 QY 361 SILSILLEEVSPELVCSMLHLCSTRLPALTVHTVTPKDGGEVCKLVGYLDRLNLEKN 420
 Db 367 SILSILLEEVSPELVCSMLHLCSTRLPALTVHTVTPKDGGEVCKLVGYLDRLNLEKN 426
 QY 421 STKQEIILAALEKGCSEFLPDYPYQKQDFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480
 Db 427 STKQEIILAALEKGCSEFLPDYPYQKQDFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 486
 QY 481 KPLLGTCKICWGPSYWCQNTETAACQNAVEHCKRHVWN 518
 Db 487 KPLLGTCKICWGPSYWCQNTETAACQNAVEHCKRHVWN 524
 RESULT 7
 ABU05203
 ID ABU05203 standard; protein; 524 AA.
 AC ABU05203;
 DT 29-JAN-2003 (first entry)
 DE Human expressed protein tag (EPT) #1869.
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 OS Homo sapiens.
 XX
 PN WO200278524-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 28-MAR-2002; 2002WO-US009671.
 XX
 PR 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 PA (ZYCO-) ZYCOS INC.
 XX
 PI Chicx RM, Tomlinson AJ, Urban RG;
 XX
 DR WPT; 2003-040607/03.
 XX
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX
 PS Example 2; SEQ ID NO 1869; 134pp; English.
 XX
 CC The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC

CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LASLGAALAGPVGLGKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKNPTVKSPLPCDIC 60
 Db 7 LASLGAALAGPVGLGKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKNPTVKSPLPCDIC 66
 QY 61 KDVTVAAGDMLKDNATEEELVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIITKGSMS 120
 Db 67 KDVTVAAGDMLKDNATEEELVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIITKGSMS 126
 QY 121 RPGEVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVWAPFMANIPILLYPQDGR 180
 Db 127 RPGEVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVWAPFMANIPILLYPQDGR 186
 QY 181 SKPQPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADICKNYIS 240
 Db 187 SKPQPKDNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADICKNYIS 246
 QY 241 QYSEIAIOMMHMPKEICALVGFCDVEKMPQTLVPAKVASKNVIIPALELVEPIKKE 300
 Db 247 QYSEIAIOMMHMPKEICALVGFCDVEKMPQTLVPAKVASKNVIIPALELVEPIKKE 306
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSECEQVWDYTG 360
 Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSECEQVWDYTG 366
 QY 361 SILSILLEEVSPELVCSMLHLCSTRLPALTVHTVTPKDGGEVCKLVGYLDRLNLEKN 420
 Db 367 SILSILLEEVSPELVCSMLHLCSTRLPALTVHTVTPKDGGEVCKLVGYLDRLNLEKN 426
 QY 421 STKQEIILAALEKGCSEFLPDYPYQKQDFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480
 Db 427 STKQEIILAALEKGCSEFLPDYPYQKQDFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 486
 QY 481 KPLLGTCKICWGPSYWCQNTETAACQNAVEHCKRHVWN 518
 Db 487 KPLLGTCKICWGPSYWCQNTETAACQNAVEHCKRHVWN 524

RESULT 8
 ABU07340
 ID ABU07340 standard; protein; 524 AA.
 XX
 AC ABU07340;
 XX
 DT 29-JAN-2003 (first entry)
 XX
 DE Human expressed protein tag (EPT) #2041.
 XX
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;

```

KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 2041; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 524 AA;
XX
XX Query Match 100.0%; Score 518; DB 6; Length 524;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1 LASLIGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTWNKPTVKSLPCDIC 60
Db 7 LASLIGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTWNKPTVKSLPCDIC 66
Qy 61 KDVVTAAGDMLKDNATBEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIRKEMS 120
Db 67 KDVVTAAGDMLKDNATBEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIRKEMS 126
Qy 121 RPGEVCSALNLCESLQKHLAEHLNKHQKLESNKIPELDMEVAVPFMANIPLLLYPQDQPR 180
Db 127 RPGEVCSALNLCESLQKHLAEHLNKHQKLESNKIPELDMEVAVPFMANIPLLLYPQDQPR 186
Qy 181 SKQPQKNGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLPGWADICKNYS 240
Db 187 SKQPQKNGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLPGWADICKNYS 246
Qy 241 QYSEIATQMMHMQPKEICALVGECDEVEKPMQTLVPKAVSKNVIPALELVEPIKKE 300
Db 247 QYSEIATQMMHMQPKEICALVGECDEVEKPMQTLVPKAVSKNVIPALELVEPIKKE 306

```

```

Qy 301 VPAKSDVYCEVCFVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEECQEVVDYTG 360
Db 307 VPAKSDVYCEVCFVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEECQEVVDYTG 366
Qy 361 SIILILEEYSPVLVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGLDRNLEKN 420
Db 367 SIILILEEYSPVLVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGLDRNLEKN 426
Qy 421 STKQEIILAALEKGCFLPDVPQKCDQFVAEYEPVLEIILVEVMDPSFVCLIKIGACPSAH 480
Db 427 STKQEIILAALEKGCFLPDVPQKCDQFVAEYEPVLEIILVEVMDPSFVCLIKIGACPSAH 486
Qy 481 KPLLGTEKCIWGPSYWCQNTETATAACNAVEHCKRHVWN 518
Db 487 KPLLGTEKCIWGPSYWCQNTETATAACNAVEHCKRHVWN 524
XX
XX RESULT 9
XX ID ABU05216 standard; protein; 524 AA.
XX
XX AC ABU05216;
XX
XX DT 29-JAN-2003 (first entry)
XX
XX DE Human expressed protein tag (EPT) #1882.
XX
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
XX OS Homo sapiens.
XX
XX PN WO200278524-A2.
XX
XX PD 10-OCT-2002.
XX
XX PF 28-MAR-2002; 2002WO-US009671.
XX
XX PR 28-MAR-2001; 2001US-0279495P.
XX 21-MAY-2001; 2001US-0292544P.
XX 08-AUG-2001; 2001US-0310801P.
XX 01-OCT-2001; 2001US-0326370P.
XX 04-DEC-2001; 2001US-0336780P.
XX 20-FEB-2002; 2002US-0358985P.
XX
XX PA (ZYCO-) ZYCOS INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1882; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 524 AA;
XX
XX Query Match 100.0%; Score 518; DB 6; Length 524;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1 LASLIGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTWNKPTVKSLPCDIC 60
Db 7 LASLIGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTWNKPTVKSLPCDIC 66
Qy 61 KDVVTAAGDMLKDNATBEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIRKEMS 120
Db 67 KDVVTAAGDMLKDNATBEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIRKEMS 126
Qy 121 RPGEVCSALNLCESLQKHLAEHLNKHQKLESNKIPELDMEVAVPFMANIPLLLYPQDQPR 180
Db 127 RPGEVCSALNLCESLQKHLAEHLNKHQKLESNKIPELDMEVAVPFMANIPLLLYPQDQPR 186
Qy 181 SKQPQKNGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLPGWADICKNYS 240
Db 187 SKQPQKNGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLPGWADICKNYS 246
Qy 241 QYSEIATQMMHMQPKEICALVGECDEVEKPMQTLVPKAVSKNVIPALELVEPIKKE 300
Db 247 QYSEIATQMMHMQPKEICALVGECDEVEKPMQTLVPKAVSKNVIPALELVEPIKKE 306

```

CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 524 AA;

```

Query Match      100.0%; Score 518; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLLGAALAGPVLGLKECTRGSAVWQNVKTSADCGAVKHCLQTVWVKPTVKSPLPCDIC 60
Db 7 LASLLGAALAGPVLGLKECTRGSAVWQNVKTSADCGAVKHCLQTVWVKPTVKSPLPCDIC 66
Qy 61 KDVVTAAGDMLKONATEEELVLEKTCWMLPKPNMSASCKEIVDSYLPVLDIIKGMS 120
Db 67 KDVVTAAGDMLKONATEEELVLEKTCWMLPKPNMSASCKEIVDSYLPVLDIIKGMS 126
Qy 121 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGPR 180
Db 127 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGPR 186
Qy 181 SKQPKDNGDVCDICQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADICKNYIS 240
Db 187 SKQPKDNGDVCDICQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADICKNYIS 246
Qy 241 QYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIIPALELVEPIKKHE 300
Db 247 QYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIIPALELVEPIKKHE 306
Qy 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDIYGS 360
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDIYGS 366
Qy 361 SILSILLEVSPELVCSMLHLCSTGRLPALTVHVTQPKDGGFCVCKKLGYLDRLNLEKN 420
Db 367 SILSILLEVSPELVCSMLHLCSTGRLPALTVHVTQPKDGGFCVCKKLGYLDRLNLEKN 426
Qy 421 STQKBIILAELKGCFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKITGACPSAH 480
Db 427 STQKBIILAELKGCFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKITGACPSAH 486
Qy 481 KPLLCTEKICWGPSYWCQNTETAACNAVEHCKRHVWN 518
Db 487 KPLLCTEKICWGPSYWCQNTETAACNAVEHCKRHVWN 524

```

RESULT 10
 ABU05202
 ID ABU05202 standard; protein; 524 AA.
 XX
 AC ABU05202;
 XX
 DT 29-JAN-2003 (first entry)
 XX
 DE Human expressed protein tag (EPT) #1868.
 XX
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 XX
 OS Homo sapiens.
 XX
 FN WO200278524-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 28-MAR-2002; 2002WO-US009671.

XX
 PR 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 PA (ZYCO-) ZYCOS INC.
 XX
 PI Chicx RM, Tomlinson AJ, Urban RG;
 XX
 DR WPI; 2003-040607/03.
 XX
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX
 PS Example 2; SEQ ID NO 1868; 134pp; English.
 XX
 CC The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, or transcription inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 524 AA;

```

Query Match      100.0%; Score 518; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLLGAALAGPVLGLKECTRGSAVWQNVKTSADCGAVKHCLQTVWVKPTVKSPLPCDIC 60
Db 7 LASLLGAALAGPVLGLKECTRGSAVWQNVKTSADCGAVKHCLQTVWVKPTVKSPLPCDIC 66
Qy 61 KDVVTAAGDMLKONATEEELVLEKTCWMLPKPNMSASCKEIVDSYLPVLDIIKGMS 120
Db 67 KDVVTAAGDMLKONATEEELVLEKTCWMLPKPNMSASCKEIVDSYLPVLDIIKGMS 126
Qy 121 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGPR 180
Db 127 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGPR 186
Qy 181 SKQPKDNGDVCDICQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADICKNYIS 240
Db 187 SKQPKDNGDVCDICQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGFGMADICKNYIS 246
Qy 241 QYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIIPALELVEPIKKHE 300
Db 247 QYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIIPALELVEPIKKHE 306
Qy 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDIYGS 360
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDIYGS 366
Qy 361 SILSILLEVSPELVCSMLHLCSTGRLPALTVHVTQPKDGGFCVCKKLGYLDRLNLEKN 420
Db 367 SILSILLEVSPELVCSMLHLCSTGRLPALTVHVTQPKDGGFCVCKKLGYLDRLNLEKN 426

```

Qy 421 STKQEIILAALEKGCFLPDYPYKQCDQFVAEYEPVLIEILVEMDPSFVCLKIGACPSAH 480
 Db 427 STKQEIILAALEKGCFLPDYPYKQCDQFVAEYEPVLIEILVEMDPSFVCLKIGACPSAH 486
 Qy 481 KPLLGTKEKCIWGPSYWCNTETAAQCNAVEHCKRHVWN 518
 Db 487 KPLLGTKEKCIWGPSYWCNTETAAQCNAVEHCKRHVWN 524

RESULT 11
 ABU05214
 ID ABU05214 standard; protein; 524 AA.
 AC ABU05214;
 XX
 DT 29-JAN-2003 (first entry)
 XX Human expressed protein tag (EPT) #1880.
 DE
 XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 XX
 OS Homo sapiens.
 XX
 XX WO200278524-A2.
 XX
 XX 10-OCT-2002.
 XX
 XX 28-MAR-2002; 2002WO-US009671.
 XX
 XX 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 XX (ZYCO-) ZYCOS INC.
 XX
 XX Chicx RM, Tomlinson AJ, Urban RG;
 XX WPI; 2003-040607/03.
 XX
 XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX
 XX Example 2; SEQ ID NO 1880; 134pp; English.
 XX
 XX The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLLGAALAGPVILGLKECTRGSAVNCQVKTASDCGAVKHCLQTVWVKPTVKSPLCDIC 60
 Db 7 LASLLGAALAGPVILGLKECTRGSAVNCQVKTASDCGAVKHCLQTVWVKPTVKSPLCDIC 66
 Qy 61 KDVTTAAGDMLKDNATEEEILVYLEKTCDWLPKPNNMSASCKEIVDSVLPVLDIILKEMS 120
 Db 67 KDVTTAAGDMLKDNATEEEILVYLEKTCDWLPKPNNMSASCKEIVDSVLPVLDIILKEMS 126
 Qy 121 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLYPQDGP 180
 Db 127 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLYPQDGP 186
 Qy 181 SKPQKNDGVDCCDCIQMVTDIQTAVRTNSTFFQALVVEHVKEBCDRGLGPGMADICKNYIS 240
 Db 187 SKPQKNDGVDCCDCIQMVTDIQTAVRTNSTFFQALVVEHVKEBCDRGLGPGMADICKNYIS 246
 Qy 241 QYSEIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVEPIKKHE 300
 Db 247 QYSEIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVEPIKKHE 306
 Qy 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPKLSSEBCQEVVDYGS 360
 Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCCKLPKLSSEBCQEVVDYGS 366
 Qy 361 SILSILLEEVSPELVCSMLHLCSTGTRLPALTAVHTVQPKDGGFCEVCKLVGYLDRLNLEKN 420
 Db 367 SILSILLEEVSPELVCSMLHLCSTGTRLPALTAVHTVQPKDGGFCEVCKLVGYLDRLNLEKN 426
 Qy 421 STKQEIILAALEKGCFLPDYPYKQCDQFVAEYEPVLIEILVEMDPSFVCLKIGACPSAH 480
 Db 427 STKQEIILAALEKGCFLPDYPYKQCDQFVAEYEPVLIEILVEMDPSFVCLKIGACPSAH 486
 Qy 481 KPLLGTKEKCIWGPSYWCNTETAAQCNAVEHCKRHVWN 518
 Db 487 KPLLGTKEKCIWGPSYWCNTETAAQCNAVEHCKRHVWN 524

RESULT 12
 ABU05215
 ID ABU05215 standard; protein; 524 AA.
 AC ABU05215;
 XX
 DT 29-JAN-2003 (first entry)
 XX Human expressed protein tag (EPT) #1881.
 DE
 XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 XX
 OS Homo sapiens.
 XX
 XX WO200278524-A2.
 XX
 XX 10-OCT-2002.
 XX
 XX 28-MAR-2002; 2002WO-US009671.
 XX
 XX 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 XX (ZYCO-) ZYCOS INC.
 XX

XX Chicx RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1881; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX lymphoma, colon cancer, gastric cancer, adenocarcinoma, melanoma,
XX leukemia or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 524 AA;

Query Match
Best Local Similarity 100.0%; Score 518; DB 6; Length 524;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LASLLGAALAGPVGLGKCTRGSAVQCNVKTASDCGAVKHCLQTVWNKPTVKSLPCDIC 60
Db 7 LASLLGAALAGPVGLGKCTRGSAVQCNVKTASDCGAVKHCLQTVWNKPTVKSLPCDIC 66
QY 61 KDVTAAAGDMLKDNATTEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIGEMS 120
Db 67 KDVTAAAGDMLKDNATTEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIGEMS 126
QY 121 RPEVCSALNLCESLOKHLAELNHQKLESNKIPELDMEVTFVAPFMANIPLLLYPDQGR 180
Db 127 RPEVCSALNLCESLOKHLAELNHQKLESNKIPELDMEVTFVAPFMANIPLLLYPDQGR 186
QY 181 SKPQPKNGDVQCQCIQWVTDIQTAVRTNSTFFQALVEHVKCECDRLGPGWADICKNYIS 240
Db 187 SKPQPKNGDVQCQCIQWVTDIQTAVRTNSTFFQALVEHVKCECDRLGPGWADICKNYIS 246
QY 241 QYSIAIQMMHMQPKICICVGVCDVEKMPQTLVPAKVASKNIPALVELVEPIKKEH 300
Db 247 QYSIAIQMMHMQPKICICVGVCDVEKMPQTLVPAKVASKNIPALVELVEPIKKEH 306
QY 301 VPAKSDVYCEVFLKVEVTKLIDNNKTEKILDAFKMCKSLPKSLSECCQVVDYGS 360
Db 307 VPAKSDVYCEVFLKVEVTKLIDNNKTEKILDAFKMCKSLPKSLSECCQVVDYGS 366
QY 361 SILSILLEEVSPELVCSMLHLCSTRLPALTVHTVQPKDGGFCEVCKLVGLVLDRLNLEKN 420
Db 367 SILSILLEEVSPELVCSMLHLCSTRLPALTVHTVQPKDGGFCEVCKLVGLVLDRLNLEKN 426
QY 421 STKQEIILAALEKCSFLPDYQKQCOQFVAEYEPVLIELVEVMDPSFVCLKIGACPSAH 480
Db 427 STKQEIILAALEKCSFLPDYQKQCOQFVAEYEPVLIELVEVMDPSFVCLKIGACPSAH 486
QY 481 KPLLGTKEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 518
Db 487 KPLLGTKEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524

RESULT 13
ABU05199
ID ABU05199 standard; protein; 524 AA.
XX
XX AC ABU05199;
XX
XX DT 29-JAN-2003 (first entry)
XX
XX DE Human expressed protein tag (EPT) #1865.
XX
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX OS Homo sapiens.
XX PN W0200278524-A2.
XX PD 10-OCT-2002.
XX
XX PF 28-MAR-2002; 2002WO-US009671.
XX
XX PR 28-MAR-2001; 2001US-0279495P.
XX PR 21-MAY-2001; 2001US-0292544P.
XX PR 08-AUG-2001; 2001US-0310801P.
XX PR 01-OCT-2001; 2001US-0326370P.
XX PR 04-DEC-2001; 2001US-0336780P.
XX PR 20-FEB-2002; 2002US-0358985P.
XX
XX PA (ZYCO-) ZYCOS INC.
XX
XX PI Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX PS Example 2; SEQ ID NO 1865; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 524 AA;

Query Match
Best Local Similarity 100.0%; Score 518; DB 6; Length 524;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LASLLGAALAGPVGLGKCTRGSAVQCNVKTASDCGAVKHCLQTVWNKPTVKSLPCDIC 60
Db 7 LASLLGAALAGPVGLGKCTRGSAVQCNVKTASDCGAVKHCLQTVWNKPTVKSLPCDIC 66
QY 61 KDVTAAAGDMLKDNATTEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIGEMS 120

Db 67 KDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKNMSASCKEIVDSYLPVLDIIGKMS 126
Qy 121 RPEVCSALNLCESLQKHLAEHLNKHQKLESNKIPELDQVTEVAVPFMANIPLLLYPDQGP 180
Db 127 RPEVCSALNLCESLQKHLAEHLNKHQKLESNKIPELDQVTEVAVPFMANIPLLLYPDQGP 186
Qy 181 SKPOKNDGVQDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRGLGPGMADICKNYIS 240
Db 187 SKPOKNDGVQDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRGLGPGMADICKNYIS 246
Qy 241 QYSEIAIQMMHMQPKEICALVGFCDVEKMPQMTLLVPAKVASKNVIPAELVEPIKKHE 300
Db 247 QYSEIAIQMMHMQPKEICALVGFCDVEKMPQMTLLVPAKVASKNVIPAELVEPIKKHE 306
Qy 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSISECQEVVDYIGS 360
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSISECQEVVDYIGS 366
Qy 361 SILSILLEEVSPELVCSMLHLCSTGTRLPALTAVHTVQPKDGGFCVCKLVGYLDRLNLEKN 420
Db 367 SILSILLEEVSPELVCSMLHLCSTGTRLPALTAVHTVQPKDGGFCVCKLVGYLDRLNLEKN 426
Qy 421 STKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 480
Db 427 STKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 486
Qy 481 KPLLGTCKCIWGPYSYWCQNTETAACQNAVEHCKRHHVN 518
Db 487 KPLLGTCKCIWGPYSYWCQNTETAACQNAVEHCKRHHVN 524

RESULT 14
ABU05212
ID ABU05212 standard; protein; 524 AA.
XX
AC ABU05212;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1878.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
XX
FI Chicx RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.

XX
PS
CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, or transcription inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 524 AA;
Query Match 100.0%; Score 518; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LASLGAALAGPVILGLKECTRGSAVWCNQVKTASDCGAVKHQCLQTVWKNKPTVKSLPCDIC 60
Db 7 LASLGAALAGPVILGLKECTRGSAVWCNQVKTASDCGAVKHQCLQTVWKNKPTVKSLPCDIC 66
Qy 61 KDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKNMSASCKEIVDSYLPVLDIIGKMS 120
Db 67 KDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKNMSASCKEIVDSYLPVLDIIGKMS 126
Qy 121 RPEVCSALNLCESLQKHLAEHLNKHQKLESNKIPELDQVTEVAVPFMANIPLLLYPDQGP 180
Db 127 RPEVCSALNLCESLQKHLAEHLNKHQKLESNKIPELDQVTEVAVPFMANIPLLLYPDQGP 186
Qy 181 SKPOKNDGVQDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRGLGPGMADICKNYIS 240
Db 187 SKPOKNDGVQDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRGLGPGMADICKNYIS 246
Qy 241 QYSEIAIQMMHMQPKEICALVGFCDVEKMPQMTLLVPAKVASKNVIPAELVEPIKKHE 300
Db 247 QYSEIAIQMMHMQPKEICALVGFCDVEKMPQMTLLVPAKVASKNVIPAELVEPIKKHE 306
Qy 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSISECQEVVDYIGS 360
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSISECQEVVDYIGS 366
Qy 361 SILSILLEEVSPELVCSMLHLCSTGTRLPALTAVHTVQPKDGGFCVCKLVGYLDRLNLEKN 420
Db 367 SILSILLEEVSPELVCSMLHLCSTGTRLPALTAVHTVQPKDGGFCVCKLVGYLDRLNLEKN 426
Qy 421 STKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 480
Db 427 STKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 486
Qy 481 KPLLGTCKCIWGPYSYWCQNTETAACQNAVEHCKRHHVN 518
Db 487 KPLLGTCKCIWGPYSYWCQNTETAACQNAVEHCKRHHVN 524

RESULT 15
ABU05213
ID ABU05213 standard; protein; 524 AA.
XX
AC ABU05213;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1879.

XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase; protease; protease inhibitor; transporter; cytoskeletal protein; receptor; transcription factor; cancer; MHC; major histocompatibility complex; myeloma; colon cancer; gastric cancer; adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX OS Homo sapiens.

XX PN WO200278524-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-US009671.

XX PR 28-MAR-2001; 2001US-0279495P.

XX PR 21-MAY-2001; 2001US-0292544P.

XX PR 08-AUG-2001; 2001US-0310801P.

XX PR 01-OCT-2001; 2001US-0326370P.

XX PR 04-DEC-2001; 2001US-0336780P.

XX PR 20-FEB-2002; 2002US-0358985P.

XX PA (ZYCO-) ZYCOS INC.

XX PI Chicx RM, Tomlinson AJ, Urban RG;

XX PF WPI; 2003-040607/03.

XX PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.

XX PS Example 2; SEQ ID NO 1879; 134pp; English.

XX CC The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukaemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLLGALAGPVLGLKECTRGSAVNCQVKTASDCGAVKHCLQTVNNKPTVKSLPCDIC 60
DB 7 LASLLGALAGPVLGLKECTRGSAVNCQVKTASDCGAVKHCLQTVNNKPTVKSLPCDIC 66

QY 61 KDVVTAAGDMLKDNATEEELVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKGENS 120
DB 67 KDVVTAAGDMLKDNATEEELVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKGENS 126

QY 121 RPEVCSALNCSLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLYPDGGR 180
DB 127 RPEVCSALNCSLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLYPDGGR 186

QY 181 SKPQKONGDVCCDICIOMVTDIQTAVRTNSTFVQALVEHVKEBCDRLPGMADICKNYIS 240
DB 187 SKPQKONGDVCCDICIOMVTDIQTAVRTNSTFVQALVEHVKEBCDRLPGMADICKNYIS 246

QY 241 QYSEIAIQMMHMQPKKEICALVGFCDVEVKEMPQMTLVPKAVSKNVIIPALELVEPIKKHE 300
DB 247 QYSEIAIQMMHMQPKKEICALVGFCDVEVKEMPQMTLVPKAVSKNVIIPALELVEPIKKHE 306

QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEECQEVVDYGS 360
DB 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEECQEVVDYGS 366

QY 361 SILSILLEEVSPELVCSMLHLCSTGLPALTVHVTOPKDGGFCEVCKLVGYLDRLNLEKN 420
DB 367 SILSILLEEVSPELVCSMLHLCSTGLPALTVHVTOPKDGGFCEVCKLVGYLDRLNLEKN 426

QY 421 STKQETLAALKEGCSFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480
DB 427 STKQETLAALKEGCSFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 486

QY 481 KPLLGTETKCIWGSPSYWCQNTETAAQCNVAHEHCKRHVWN 518
DB 487 KPLLGTETKCIWGSPSYWCQNTETAAQCNVAHEHCKRHVWN 524

Search completed: May 5, 2004, 13:45:39
Job time : 54.1919 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:28:28 ; Search time 15.405 Seconds
(without alignments)
1630.129 Million cell updates/sec

Title: US-09-743-684A-1_COPY_7_524

Perfect score: 518
Sequence: 1 LASLLGAALAGPVLGLKECT.....NTETAQCNAVEHCKRHWN 518

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	518	100.0	524	4	US-09-352-548-1
2	224	43.2	523	1	US-08-100-247-2
3	224	43.2	523	1	US-08-232-513A-3
4	224	43.2	523	4	US-08-756-031-2
5	187	36.1	523	1	US-08-483-146A-2
6	187	36.1	523	1	US-08-484-594A-2
7	187	36.1	523	4	US-09-076-258A-2
8	81	15.6	81	4	US-09-352-548-2
9	80	15.4	80	1	US-08-100-247-3
10	80	15.4	80	1	US-08-483-146A-3
11	80	15.4	80	1	US-08-232-513A-4
12	80	15.4	80	1	US-08-484-594A-3
13	80	15.4	80	4	US-09-076-258A-3
14	80	15.4	80	4	US-08-756-031-3
15	75	14.5	80	2	US-08-584-671-15
16	75	14.5	80	3	US-09-027-376-15
17	69	13.3	80	3	US-09-094-192-15
18	61	11.8	61	2	US-08-584-671-13
19	61	11.8	61	3	US-09-027-376-13
20	61	11.8	61	3	US-09-094-192-13
21	61	11.8	69	3	US-09-268-070-2
22	22	4.2	22	1	US-08-100-247-1
23	22	4.2	22	1	US-08-483-146A-1
24	22	4.2	22	1	US-08-483-146A-11
25	22	4.2	22	1	US-08-232-513A-1
26	22	4.2	22	1	US-08-232-513A-20
27	22	4.2	22	1	US-08-484-594A-1

28	22	4.2	22	1	US-08-484-594A-11	Sequence 11, Appl
29	22	4.2	22	3	US-09-231-159-1	Sequence 1, Appl
30	22	4.2	22	3	US-08-611-307-1	Sequence 1, Appl
31	22	4.2	22	4	US-09-148-030-1	Sequence 1, Appl
32	22	4.2	22	4	US-09-076-258A-1	Sequence 1, Appl
33	22	4.2	22	4	US-09-076-258A-11	Sequence 11, Appl
34	22	4.2	22	4	US-08-756-031-1	Sequence 1, Appl
35	18	3.5	18	4	US-09-352-548-17	Sequence 17, Appl
36	17	3.3	17	4	US-09-352-548-14	Sequence 14, Appl
37	17	3.3	17	4	US-09-352-548-15	Sequence 15, Appl
38	17	3.3	18	1	US-08-100-247-5	Sequence 5, Appl
39	17	3.3	18	1	US-08-483-146A-5	Sequence 5, Appl
40	17	3.3	18	1	US-08-232-513A-6	Sequence 6, Appl
41	17	3.3	18	1	US-08-484-594A-5	Sequence 5, Appl
42	17	3.3	18	3	US-09-231-159-20	Sequence 20, Appl
43	17	3.3	18	3	US-08-611-307-20	Sequence 20, Appl
44	17	3.3	18	4	US-09-148-030-2	Sequence 2, Appl
45	17	3.3	18	4	US-09-076-258A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-352-548-1
; Sequence 1, Application US/09352548
; Patent No. 6500431
; GENERAL INFORMATION:
; APPLICANT: Gill, Parkash S.
; APPLICANT: Parkash S. Gill, M.D., Inc.
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
; FILE REFERENCE: 017986-000410US
; CURRENT APPLICATION NUMBER: US/09/352.548
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: US 60/092.647
; EARLIER FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: prosaposin
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (195)..(275)
; OTHER INFORMATION: Saposin B
US-09-352-548-1

Query Match	100.0%;	Score 518;	DB 4;	Length 524;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 518;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	LASLLGAALAGPVLGLKECTRGSAVQCQNVKTASDCGAVKHCLQTVMNKPTVKSLPCDIC	60	
Db	7	LASLLGAALAGPVLGLKECTRGSAVQCQNVKTASDCGAVKHCLQTVMNKPTVKSLPCDIC	66	
Qy	61	KDVVTAAGDMLKDNATEEEILVYLEKCDMLPKFNMSASCKEIVDSYLPVILDIKIGMS	120	
Db	67	KDVVTAAGDMLKDNATEEEILVYLEKCDMLPKFNMSASCKEIVDSYLPVILDIKIGMS	126	
Qy	121	RPGEVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVVAFPMANIPLLLYPDQDGR	180	
Db	127	RPGEVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVVAFPMANIPLLLYPDQDGR	186	
Qy	181	SKPQKDNQGVQCDCIQMVTDIQTAVRTNSTFFQALVEHVKEECDRLPGMGADICKNYIS	240	
Db	187	SKPQKDNQGVQCDCIQMVTDIQTAVRTNSTFFQALVEHVKEECDRLPGMGADICKNYIS	246	
Qy	241	QYSEIATQMMHMQPKEICALVGFCDEVKEMPQMTLVPKVASKNVIPAELVEPIKKHE	300	
Db	247	QYSEIATQMMHMQPKEICALVGFCDEVKEMPQMTLVPKVASKNVIPAELVEPIKKHE	306	

QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMKSKLPKSLSECEQEVVDITYGS 360
DB 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMKSKLPKSLSECEQEVVDITYGS 366
QY 361 SILSILILEEVSPELVCSMLHLCSTGTRLPALTVHTVTPKDGCGFCEVCKKLVGLVDRNLEKN 420
DB 367 SILSILILEEVSPELVCSMLHLCSTGTRLPALTVHTVTPKDGCGFCEVCKKLVGLVDRNLEKN 426
QY 421 STKQEIILAALEKSGFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480
DB 427 STKQEIILAALEKSGFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 486
QY 481 KPLLGTCKIWGSPSYWCQNTETAACQNAVEHCKRHWVN 518
DB 487 KPLLGTCKIWGSPSYWCQNTETAACQNAVEHCKRHWVN 524

RESULT 2
US-08-100-247-2
; Sequence 2, Application US/08100247
; Patent No. 5571787
; GENERAL INFORMATION:
; APPLICANT: O'BRIEN, JOHN S.
; APPLICANT: KISHIMOTO, YASUO
; TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/100.247
; FILING DATE: 19930730
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: O'Brien.002A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; IMMEDIATE SOURCE:
; CLONE: PROSAPOSIN
US-08-100-247-2

Query Match 43.2%; Score 224; DB 1; Length 523;
Best Local Similarity 100.0%; Pred. No. 1e-207;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 295 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMKSKLPKSLSECEQEV 354
DB 300 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMKSKLPKSLSECEQEV 359
QY 355 VDTYGSILSILLEEVSPELVCSMLHLCSTGTRLPALTVHTVTPKDGCGFCEVCKKLVGLYLD 414

DB 360 VDTYGSILSILLEEVSPELVCSMLHLCSTGTRLPALTVHTVTPKDGCGFCEVCKKLVGLYLD 419
QY 415 RNLEKNTKQEIILAALEKSGFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 474
DB 420 RNLEKNTKQEIILAALEKSGFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 479
QY 475 ACPSAHKPLLGTEKCIWGSPSYWCQNTETAACQNAVEHCKRHWVN 518
DB 480 ACPSAHKPLLGTEKCIWGSPSYWCQNTETAACQNAVEHCKRHWVN 523

RESULT 3
US-08-232-513A-3
; Sequence 3, Application US/08232513A
; Patent No. 5700909
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides
; TITLE OF INVENTION: as Therapeutic Agents
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.513A
; FILING DATE: 21-APR-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/100.247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 1643
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..523
; OTHER INFORMATION: /label= Hum_prosaposin
US-08-232-513A-3

Query Match 43.2%; Score 224; DB 1; Length 523;
Best Local Similarity 100.0%; Pred. No. 1e-207;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 295 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMKSKLPKSLSECEQEV 354
DB 300 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMKSKLPKSLSECEQEV 359
QY 355 VDTYGSILSILLEEVSPELVCSMLHLCSTGTRLPALTVHTVTPKDGCGFCEVCKKLVGLYLD 414
DB 360 VDTYGSILSILLEEVSPELVCSMLHLCSTGTRLPALTVHTVTPKDGCGFCEVCKKLVGLYLD 419
QY 415 RNLEKNTKQEIILAALEKSGFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 474

Db 420 RNLEKNTKQBILAALEKGCFLPDPYQKQCDQFVAEYEPVLBIILVEVMDPSFVCLKIG 479

Qy 475 ACPSAHKPLLGTEKCIWGPSYWCNTETATAACNAVEHCKRHVMN 518
|||||

Db 480 ACPSAHKPLLGTEKCIWGPSYWCNTETATAACNAVEHCKRHVMN 523
|||||

RESULT 4

US-08-756-031-2

; Sequence 2, Application US/08756031

; Patent No. 6590074

; GENERAL INFORMATION:

; APPLICANT: O'BRIEN, JOHN S.

; APPLICANT: KISHIMOTO, YASUO

; TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR

; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR

; CITY: NEWPORT BEACH

; STATE: CA

; COUNTRY: USA

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/756,031

; FILING DATE: 26-NOV-1996

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/100,247

; FILING DATE: 30-JUL-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Israelsen, Ned A.

; REGISTRATION NUMBER: 29,655

; REFERENCE/DOCKET NUMBER: O'BRIEN.002A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-235-8550

; TELEFAX: 619-235-0176

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 523 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

; IMMEDIATE SOURCE:

; CLONE: PROSAPOSIN

US-08-756-031-2

Query Match 43.2%; Score 224; DB 4; Length 523;

Best Local Similarity 100.0%; Pred. No. 1e-207;

Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 295 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKVCCKLPKSLSEECQEV 354
|||||

Db 300 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKVCCKLPKSLSEECQEV 359
|||||

Qy 355 VDTYGSSTLSLLBEVSPPELVCSMLHLCSTGRFLPALTTHVTQPKDGGFCVCKLVGYLD 414
|||||

Db 360 VDTYGSSTLSLLBEVSPPELVCSMLHLCSTGRFLPALTTHVTQPKDGGFCVCKLVGYLD 419
|||||

Qy 415 RNLEKNTKQBILAALEKGCFLPDPYQKQCDQFVAEYEPVLBIILVEVMDPSFVCLKIG 474
|||||

Db 420 RNLEKNTKQBILAALEKGCFLPDPYQKQCDQFVAEYEPVLBIILVEVMDPSFVCLKIG 479
|||||

Qy 475 ACPSAHKPLLGTEKCIWGPSYWCNTETATAACNAVEHCKRHVMN 518
|||||

Db 480 ACPSAHKPLLGTEKCIWGPSYWCNTETATAACNAVEHCKRHVMN 523
|||||

RESULT 5

US-08-483-146A-2

; Sequence 2, Application US/08483146A

; Patent No. 5696080

; GENERAL INFORMATION:

; APPLICANT: O'Brien, John S.

; APPLICANT: Kishimoto, Yasuo

; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS

; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED THEREFROM

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson and Bear

; STREET: 620 Newport Center Blvd. 16th Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: USA

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/483,146A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Israelsen, Ned A.

; REGISTRATION NUMBER: 29,655

; REFERENCE/DOCKET NUMBER: MYELOS.002DV1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-235-8550

; TELEFAX: 619-235-0176

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 523 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: N-terminal

US-08-483-146A-2

Query Match 36.1%; Score 187; DB 1; Length 523;

Best Local Similarity 100.0%; Pred. No. 5.2e-172;

Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLLGAALAGPVLGLKECTRGSAVWCNVKATSDCGAVKHCLQTVWNKPTVKSLPCDIC 60
|||||

Db 7 LASLLGAALAGPVLGLKECTRGSAVWCNVKATSDCGAVKHCLQTVWNKPTVKSLPCDIC 66
|||||

Qy 61 KDVVTAAGMDLKNATATEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKGEWS 120
|||||

Db 67 KDVVTAAGMDLKNATATEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKGEWS 126
|||||

Qy 121 RPEGVCSALMLCESLQKHAEHLNKHQKLESNKIPELDMTEVVAFFMANIPILLIYPQDGP 180
|||||

Db 127 RPEGVCSALMLCESLQKHAEHLNKHQKLESNKIPELDMTEVVAFFMANIPILLIYPQDGP 186
|||||

Qy 181 SKPOPKD 187
|||||

Db 187 SKPOPKD 193
|||||

RESULT 6

US-08-484-594A-2
; Sequence 2, Application US/08484594A
; Patent No. 5714459
; GENERAL INFORMATION:

APPLICANT: O'Brien, John S.
APPLICANT: Kishimoto, Yasuo
TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUROTROPHIC PEPTIDES
TITLE OF INVENTION: DERIVED THEREFROM
NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA

ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,594A
FILING DATE: 07-JUN-1995

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/100,247

FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A

REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.002DV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: Peptide
FRAGMENT TYPE: N-terminal

US-08-484-594A-2

Query Match 36.1%; Score 187; DB 1; Length 523;
Best Local Similarity 100.0%; Pred. No. 5.2e-172;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSPLPCDIC	60
Db	7	LASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSPLPCDIC	66
QY	61	KDVVTAAGDMLKONATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKEMS	120
Db	67	KDVVTAAGDMLKONATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKEMS	126
QY	121	RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGGR	180
Db	127	RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGGR	186
QY	181	SKPQPKD	187
Db	187	SKPQPKD	193

RESULT 7

US-09-076-258A-2
; Sequence 2, Application US/09076258A
; Patent No. 6559124
; GENERAL INFORMATION:

APPLICANT: O'Brien, John S.
APPLICANT: Kishimoto, Yasuo
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED
TITLE OF INVENTION: THEREFROM
NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Blvd. 15th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA

ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/076,258A
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/958,970
FILING DATE: 28-OCT-97

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/483,146
FILING DATE: 07-JUN-1995

APPLICATION NUMBER: 08/100,247
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.2DVIC2

TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: Peptide
FRAGMENT TYPE: N-terminal

US-09-076-258A-2

Query Match 36.1%; Score 187; DB 4; Length 523;
Best Local Similarity 100.0%; Pred. No. 5.2e-172;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSPLPCDIC	60
Db	7	LASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSPLPCDIC	66
QY	61	KDVVTAAGDMLKONATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKEMS	120
Db	67	KDVVTAAGDMLKONATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKEMS	126
QY	121	RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGGR	180
Db	127	RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGGR	186
QY	181	SKPQPKD	187
Db	187	SKPQPKD	193

RESULT 8

US-09-352-548-2
; Sequence 2, Application US/09352548
; Patent No. 6500431

GENERAL INFORMATION:
APPLICANT: Gill, Parkash S.
TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
FILE REFERENCE: 017986-000410US
CURRENT APPLICATION NUMBER: US/09/352,548
CURRENT FILING DATE: 1999-07-12
EARLIER APPLICATION NUMBER: US 60/092,647
EARLIER FILING DATE: 1998-07-13
NUMBER OF SEQ ID NOS: 59
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 81
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Saposin B
US-09-352-548-2

Query Match 15.6%; Score 81; DB 4; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.8e-70;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 189 GDVQCDCIQWYDTDTAVRTNSTFVQALVHVHKEECDRLGFGMADICKNYISQYSEIAIQ 248
Db 1 GDVQCDCIQWYDTDTAVRTNSTFVQALVHVHKEECDRLGFGMADICKNYISQYSEIAIQ 60
Qy 249 MMMHMQKEICALVGFCDVEK 269
Db 61 MMMHMQKEICALVGFCDVEK 81

RESULT 9
US-08-100-247-3
Sequence 3, Application US/08100247
Patent No. 5571787
GENERAL INFORMATION:
APPLICANT: O'BRIEN, JOHN S.
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/100,247
FILING DATE: 19930730
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: O'BRIEN.002A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO

FRAGMENT TYPE: internal
IMMEDIATE SOURCE:
CLONE: SAPOSIN C
US-08-100-247-3
Query Match 15.4%; Score 80; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.7e-69;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 305 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSECOEVDVDTYGSLS 364
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSECOEVDVDTYGSLS 60
Qy 365 ILLEEVSPELVCSMLHLCSG 384
Db 61 ILLEEVSPELVCSMLHLCSG 80
RESULT 10
US-08-483-146A-3
Sequence 3, Application US/08483146A
Patent No. 5696080
GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED
TITLE OF INVENTION: THEREFROM
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Blvd. 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,146A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.002DV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-483-146A-3
Query Match 15.4%; Score 80; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.7e-69;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 305 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSECOEVDVDTYGSLS 364
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSECOEVDVDTYGSLS 60

QY 365 ILLEVPSPVLCVSMHLCSG 384
Db 61 ILLEVPSPVLCVSMHLCSG 80

RESULT 11

US-08-232-513A-4
; Sequence 4, Application US/08232513A
; Patent No. 5700909
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; TITLE OF INVENTION: Proasoposin and Cytokine-Derived Peptides
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,513A
; FILING DATE: 21-APR-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/100,247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 1643
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-8949
; TELEFAX: (619) 535-9001
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..80
; OTHER INFORMATION: /label= Saposin_C
US-08-232-513A-4

Query Match 15.4%; Score 80; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.7e-69;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 305 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDITYGSSILS 364
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDITYGSSILS 60
QY 365 ILLEVPSPVLCVSMHLCSG 384
Db 61 ILLEVPSPVLCVSMHLCSG 80

RESULT 12

US-08-484-594A-3
; Sequence 3, Application US/08484594A
; Patent No. 5714459
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; APPLICANT: Kishimoto, Yasuo

; TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUROTROPHIC PEPTIDES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,594A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/100,247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: MYELOS.002DV2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-484-594A-3
Query Match 15.4%; Score 80; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.7e-69;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 305 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDITYGSSILS 364
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDITYGSSILS 60
QY 365 ILLEVPSPVLCVSMHLCSG 384
Db 61 ILLEVPSPVLCVSMHLCSG 80
RESULT 13
US-09-076-258A-3
; Sequence 3, Application US/09076258A
; Patent No. 6559124
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; APPLICANT: Kishimoto, Yasuo
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Blvd. 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/076,258A
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/958,970
FILING DATE: 28-OCT-97

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/483,146
FILING DATE: 07-JUN-1995

APPLICATION NUMBER: 08/100,247
FILING DATE: 30-JUL-1993

ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.

REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.2DV1C2

TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 80 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

US-09-076-258A-3

Query Match 15.4%; Score 80; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.7e-69;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDVTYGSILS 364
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDVTYGSILS 60

Qy 365 ILLEEVSPELVCSMLHLCSG 384
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 14

US-08-756-031-3
Sequence 3, Application US/08756031
Patent No. 6590074

GENERAL INFORMATION:

APPLICANT: O'BRIEN, JOHN S.

APPLICANT: KISHIMOTO, YASUO

TITLE OF INVENTION: PROSAPIN AS A NEUROTROPIC FACTOR

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR

STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR

CITY: NEWPORT BEACH

STATE: CA

COUNTRY: USA

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/756,031

FILING DATE: 26-NOV-1996

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/100,247

FILING DATE: 30-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned A.

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: O'BRIEN.002A

TELEPHONE: 619-235-8550

TELEFAX: 619-235-0176

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 80 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

IMMEDIATE SOURCE:

CLONE: SAPOSIN C

US-08-756-031-3

Query Match 15.4%; Score 80; DB 4; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.7e-69;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDVTYGSILS 364
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDVTYGSILS 60

Qy 365 ILLEEVSPELVCSMLHLCSG 384
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 15

US-08-584-671-15
Sequence 15, Application US/08584671

Patent No. 5910568

GENERAL INFORMATION:

APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,

APPLICANT: CRAMER, PALMER

TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM

TITLE OF INVENTION: TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE

TITLE OF INVENTION: TO ENHANCE OR DECREASE POTENTIAL FERTILITY

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA

ADDRESSEE: STATE UNIVERSITY

STREET: 113 TECHNOLOGY CENTER

CITY: UNIVERSITY PARK

STATE: PENNSYLVANIA

COUNTRY: UNITED STATES OF AMERICA

ZIP: 16802-7000

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: NEC-286

OPERATING SYSTEM: DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/584,671

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MONAHAN, THOMAS J

REGISTRATION NUMBER: 29835

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: 814-865-6277

TELEFAX: 814-865-3591

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 80

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:33:49 ; Search time 41.7582 Seconds
(without alignments)
3438.471 Million cell updates/sec

Title: US-09-743-684A-1_COPY_7_524

Perfect score: 518

Sequence: 1 LASLLGAALAGPVLGLKECT.....NTETAACQNAVEHCKRHVWN 518

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1138120 seqs, 277189581 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/PCT PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT NEW PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06 NEW PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06 PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07 NEW PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08 NEW PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08 PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09 NEW PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10 NEW PUB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60 NEW PUB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60 PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	518	100.0	524	9	US-09-870-759-60
2	518	100.0	524	10	US-09-751-708A-60
3	518	100.0	524	12	US-10-267-502-386
4	278	53.7	527	9	US-09-870-759-61
5	278	53.7	527	10	US-09-751-708A-61
6	278	53.7	527	14	US-10-060-036-73
7	254	49.0	479	10	US-09-978-418-40
8	209	40.3	209	14	US-10-043-487-340
9	187	36.1	523	9	US-09-767-007A-2
10	80	15.4	80	9	US-09-767-007A-3
11	80	15.4	80	9	US-09-753-126-3
12	80	15.4	80	15	US-10-330-697-3
13	80	15.4	592	9	US-09-753-126-4
14	80	15.4	592	15	US-10-330-697-4
15	40	7.7	40	10	US-09-780-438A-1

16	38	7.3	38	10	US-09-780-438A-2	Sequence 2, Appli
17	156	12	156	12	US-09-925-298-644	Sequence 644, App
18	24	4.6	156	14	US-10-102-806-644	Sequence 644, App
19	22	4.2	22	9	US-09-767-007A-1	Sequence 1, Appli
20	22	4.2	22	9	US-09-767-007A-11	Sequence 11, Appli
21	22	4.2	22	9	US-09-957-143-1	Sequence 1, Appli
22	20	3.9	554	14	US-10-205-194-176	Sequence 176, App
23	17	3.3	18	9	US-09-767-007A-5	Sequence 5, Appli
24	17	3.3	18	9	US-09-957-143-2	Sequence 2, Appli
25	15	2.9	15	9	US-09-802-617-2	Sequence 2, Appli
26	15	2.9	15	14	US-10-293-819-2	Sequence 2, Appli
27	12	2.3	12	9	US-09-802-617-1	Sequence 1, Appli
28	12	2.3	12	9	US-09-957-143-3	Sequence 3, Appli
29	12	2.3	12	14	US-10-293-819-1	Sequence 1, Appli
30	11	2.1	11	10	US-09-791-393-182	Sequence 182, App
31	11	2.1	11	10	US-09-791-389-182	Sequence 182, App
32	11	2.1	11	14	US-10-281-478-126	Sequence 126, App
33	10	1.9	15	14	US-10-281-478-74	Sequence 74, Appli
34	10	1.9	20	14	US-10-281-478-77	Sequence 77, Appli
35	9	1.7	14	14	US-10-281-478-75	Sequence 129, App
36	9	1.7	17	14	US-10-281-478-129	Sequence 4623, Ap
37	8	1.5	286	15	US-10-108-260A-4623	Sequence 279814, A
38	8	1.5	354	12	US-10-424-599-279814	Sequence 46329, A
39	8	1.5	361	12	US-10-425-114-46329	Sequence 72368, A
40	8	1.5	363	12	US-10-425-114-72368	Sequence 10328, A
41	8	1.5	655	9	US-09-815-242-10328	Sequence 56714, A
42	8	1.5	655	12	US-10-282-122A-56714	Sequence 34, Appli
43	8	1.5	1896	12	US-10-312-352-34	Sequence 35398, A
44	7	1.4	46	9	US-09-864-761-35398	Sequence 211534, A
45	7	1.4	50	12	US-10-424-599-211534	

ALIGNMENTS

RESULT 1
US-09-870-759-60
; Sequence 60, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 60
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-60

Query Match	100.0%;	Score 518;	DB 9;	Length 524;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 518;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	LASLLGAALAGPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWKNKPTVKSLPCDIC	60	
Db	7	LASLLGAALAGPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWKNKPTVKSLPCDIC	66	
Qy	61	KDVVTAAGDMLKDNATEEILVLEKTCMDLKPKNMSASCKEIVDSYLPVILDIKIGMS	120	
Db	67	KDVVTAAGDMLKDNATEEILVLEKTCMDLKPKNMSASCKEIVDSYLPVILDIKIGMS	126	
Qy	121	RPGEVCSALNLCESLQKHAEHNLHQLKESNKPISPELDMTEVAVPFMANIPLLLYPDQDPR	180	
Db	127	RPGEVCSALNLCESLQKHAEHNLHQLKESNKPISPELDMTEVAVPFMANIPLLLYPDQDPR	186	
Qy	181	SKPQKONGDVCDQCIQMTVDIQTAVRTNSTFFQVALVHVKECDRLGPGMADICKNYS	240	

Db 187 SKPQKONGDVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPMADICKNYIS 246
 QY 241 QYSEIAIQMMHMQPKEICALVGFCDEVKEMPQTLVPAKASKNVI PALELVEPIKKHE 300
 Db 247 QYSEIAIQMMHMQPKEICALVGFCDEVKEMPQTLVPAKASKNVI PALELVEPIKKHE 306
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSECEQVVDITYGS 360
 Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSECEQVVDITYGS 366
 QY 361 SILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCVCKKLVGLDRNLEKN 420
 Db 367 SILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCVCKKLVGLDRNLEKN 426
 QY 421 STKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480
 Db 427 STKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 486
 QY 481 KPLLGTKEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 518
 Db 487 KPLLGTKEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524
 RESULT 2
 US-09-751-708A-60
 ; Sequence 60, Application US/09751708A
 ; Publication No. US20030157113A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TERMAN, David S
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
 ; FILE REFERENCE: 751708
 ; CURRENT APPLICATION NUMBER: US/09/751,708A
 ; CURRENT FILING DATE: 2002-10-15
 ; PRIOR APPLICATION NUMBER: US 60/173,371
 ; PRIOR FILING DATE: 1999-12-28
 ; NUMBER OF SEQ ID NOS: 166
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 60
 ; LENGTH: 524
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-751-708A-60

Query Match 100.0%; Score 518; DB 10; Length 524;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKSLLPCDIC 60
 Db 7 LASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKSLLPCDIC 66
 QY 61 KDVTAAAGDMLKDNATEEELVLEKTCWLPKPNMSASCKEIVDSYLPVILDIKGENS 120
 Db 67 KDVTAAAGDMLKDNATEEELVLEKTCWLPKPNMSASCKEIVDSYLPVILDIKGENS 126
 QY 121 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFPMANIPLLLYPQDQPR 180
 Db 127 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFPMANIPLLLYPQDQPR 186
 QY 181 SKPQKONGDVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPMADICKNYIS 240
 Db 187 SKPQKONGDVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPMADICKNYIS 246
 QY 241 QYSEIAIQMMHMQPKEICALVGFCDEVKEMPQTLVPAKASKNVI PALELVEPIKKHE 300
 Db 247 QYSEIAIQMMHMQPKEICALVGFCDEVKEMPQTLVPAKASKNVI PALELVEPIKKHE 306
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSECEQVVDITYGS 360
 Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSECEQVVDITYGS 366
 QY 361 SILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCVCKKLVGLDRNLEKN 420
 Db 367 SILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCVCKKLVGLDRNLEKN 426
 QY 421 STKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480
 Db 427 STKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 486
 QY 481 KPLLGTKEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 518
 Db 487 KPLLGTKEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524
 RESULT 4
 US-09-751-708A-60

Db 367 SILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCVCKKLVGLDRNLEKN 426
 QY 421 STKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480
 Db 427 STKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 486
 QY 481 KPLLGTKEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 518
 Db 487 KPLLGTKEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524
 RESULT 3
 US-10-267-502-386
 ; Sequence 386, Application US/10267502
 ; Publication No. US20040071700A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kim, Jaeseob
 ; APPLICANT: Galant, Ron
 ; TITLE OF INVENTION: Obesity Linked Genes
 ; FILE REFERENCE: LSD-07416
 ; CURRENT APPLICATION NUMBER: US/10/267,502
 ; CURRENT FILING DATE: 2003-01-27
 ; NUMBER OF SEQ ID NOS: 439
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 386
 ; LENGTH: 524
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-267-502-386

Query Match 100.0%; Score 518; DB 12; Length 524;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKSLLPCDIC 60
 Db 7 LASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKSLLPCDIC 66
 QY 61 KDVTAAAGDMLKDNATEEELVLEKTCWLPKPNMSASCKEIVDSYLPVILDIKGENS 120
 Db 67 KDVTAAAGDMLKDNATEEELVLEKTCWLPKPNMSASCKEIVDSYLPVILDIKGENS 126
 QY 121 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFPMANIPLLLYPQDQPR 180
 Db 127 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFPMANIPLLLYPQDQPR 186
 QY 181 SKPQKONGDVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPMADICKNYIS 240
 Db 187 SKPQKONGDVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPMADICKNYIS 246
 QY 241 QYSEIAIQMMHMQPKEICALVGFCDEVKEMPQTLVPAKASKNVI PALELVEPIKKHE 300
 Db 247 QYSEIAIQMMHMQPKEICALVGFCDEVKEMPQTLVPAKASKNVI PALELVEPIKKHE 306
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSECEQVVDITYGS 360
 Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSECEQVVDITYGS 366
 QY 361 SILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCVCKKLVGLDRNLEKN 420
 Db 367 SILSILLEEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCVCKKLVGLDRNLEKN 426
 QY 421 STKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480
 Db 427 STKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 486
 QY 481 KPLLGTKEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 518
 Db 487 KPLLGTKEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524
 RESULT 4
 US-09-870-759-61

```
; Sequence 61, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-61

Query Match      53.7%; Score 278; DB 9; Length 527;
Best Local Similarity 99.4%; Pred. No. 1.9e-263;
Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 LASLLGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTVWKNKPTVKSILPCDIC 60
Db 7 LASLLGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTVWKNKPTVKSILPCDIC 66
Qy 61 KDVVTAAGDMLKDNATEEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKIGEMS 120
Db 67 KDVVTAAGDMLKDNATEEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKIGEMS 126
Qy 121 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLYPQDQPR 180
Db 127 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLYPQDQPR 186
Qy 181 SKQPKDNGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRILGPGMADICKNYIS 240
Db 187 SKQPKDNGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRILGPGMADICKNYIS 246
Qy 241 QYSEIAIQMMHMM---OPKEICALVGFCDVEKEMPMOTLVPKAVSKNVIIPALELVEPIK 297
Db 247 QYSEIAIQMMHMMQDQPKKEICALVGFCDVEKEMPMOTLVPKAVSKNVIIPALELVEPIK 306
Qy 298 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDT 357
Db 307 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDT 366
Qy 358 YGSSILSILLLEVSPELVCSMLHLCSTGTRLPALTVHTVTPKDGGFCEVCKKLGYLDRL 417
Db 367 YGSSILSILLLEVSPELVCSMLHLCSTGTRLPALTVHTVTPKDGGFCEVCKKLGYLDRL 426
Qy 418 EKNSTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIGACP 477
Db 427 EKNSTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIGACP 486

RESULT 5
US-09-751-708A-61
; Sequence 61, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 61
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-751-708A-61

Query Match      53.7%; Score 278; DB 10; Length 527;
Best Local Similarity 99.4%; Pred. No. 1.9e-263;
Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 LASLLGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTVWKNKPTVKSILPCDIC 60
Db 7 LASLLGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTVWKNKPTVKSILPCDIC 66
Qy 61 KDVVTAAGDMLKDNATEEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKIGEMS 120
Db 67 KDVVTAAGDMLKDNATEEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKIGEMS 126
Qy 121 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLYPQDQPR 180
Db 127 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLYPQDQPR 186
Qy 181 SKQPKDNGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRILGPGMADICKNYIS 240
Db 187 SKQPKDNGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRILGPGMADICKNYIS 246
Qy 241 QYSEIAIQMMHMM---OPKEICALVGFCDVEKEMPMOTLVPKAVSKNVIIPALELVEPIK 297
Db 247 QYSEIAIQMMHMMQDQPKKEICALVGFCDVEKEMPMOTLVPKAVSKNVIIPALELVEPIK 306
Qy 298 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDT 357
Db 307 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDT 366
Qy 358 YGSSILSILLLEVSPELVCSMLHLCSTGTRLPALTVHTVTPKDGGFCEVCKKLGYLDRL 417
Db 367 YGSSILSILLLEVSPELVCSMLHLCSTGTRLPALTVHTVTPKDGGFCEVCKKLGYLDRL 426
Qy 418 EKNSTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIGACP 477
Db 427 EKNSTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIGACP 486

RESULT 6
US-10-060-036-73
; Sequence 73, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-036-73

Query Match      53.7%; Score 278; DB 14; Length 527;
Best Local Similarity 99.4%; Pred. No. 1.9e-263;
```

Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 LASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVNNKPTVKSLPCDLC 60
Db 7 LASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVNNKPTVKSLPCDLC 66

QY 61 KDVTAAAGDMLKONATEEEILVYLEKTCWLPKPNMSASCKEIVDSYLPVILDIKIGEMS 120
Db 67 KDVTAAAGDMLKONATEEEILVYLEKTCWLPKPNMSASCKEIVDSYLPVILDIKIGEMS 126

QY 121 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLYPDQGP 180
Db 127 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLYPDQGP 186

QY 181 SKQPKDNGDVCDCCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 240
Db 187 SKQPKDNGDVCDCCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 246

QY 241 QYSEIAIQMMHMQ 254
Db 247 QYSEIAIQMMHMQ 260

RESULT 8
US-10-043-487-340
; Sequence 340, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEGRAN
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 340
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Shigella Flexneri
US-10-043-487-340

Query Match 40.3%; Score 209; DB 14; Length 209;
Best Local Similarity 100.0%; Pred. No. 3.4e-196;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 OKLESNKIPELDMTEVAVPFMANIPLLLYPDQGPGRSKPOKNGDVCQDCIQMTVDIQT 204
Db 1 OKLESNKIPELDMTEVAVPFMANIPLLLYPDQGPGRSKPOKNGDVCQDCIQMTVDIQT 60

QY 205 AVRTNSTFVQALVEHVKEECDRLGPGMADICKNYISOYSEIAIQMMHMQPKETCALVGF 264
Db 61 AVRTNSTFVQALVEHVKEECDRLGPGMADICKNYISOYSEIAIQMMHMQPKETCALVGF 120

QY 265 CDEVKEMPQTLVPAKVASKNVIPALELVEPIKKEHVEPAKSDVYCEVCEFLVKEVTKLID 324
Db 121 CDEVKEMPQTLVPAKVASKNVIPALELVEPIKKEHVEPAKSDVYCEVCEFLVKEVTKLID 180

QY 325 NNKTEKILDAFDKMSCKLPKSLSEECQE 353
Db 181 NNKTEKILDAFDKMSCKLPKSLSEECQE 209

RESULT 9
US-09-767-007A-2
; Sequence 2, Application US/09767007A
; Patent No. US20020077275A1
; GENERAL INFORMATION:
; APPLICANT: John S. O'Brien
; APPLICANT: Yasuo Kishimoto
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
; FILE REFERENCE: SAPOSIN C AND NEUROTROPIC PEPTIDES DERIVED THEREFROM
; CURRENT APPLICATION NUMBER: US/09/767,007A
; CURRENT FILING DATE: 2001-01-22

Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 LASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVNNKPTVKSLPCDLC 60
Db 7 LASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVNNKPTVKSLPCDLC 66

QY 61 KDVTAAAGDMLKONATEEEILVYLEKTCWLPKPNMSASCKEIVDSYLPVILDIKIGEMS 120
Db 67 KDVTAAAGDMLKONATEEEILVYLEKTCWLPKPNMSASCKEIVDSYLPVILDIKIGEMS 126

QY 121 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLYPDQGP 180
Db 127 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLYPDQGP 186

QY 181 SKQPKDNGDVCDCCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 240
Db 187 SKQPKDNGDVCDCCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 246

QY 241 QYSEIAIQMMHMQ 254
Db 247 QYSEIAIQMMHMQ 260

RESULT 7
US-09-978-418-40
; Sequence 40, Application US/09978418
; Publication No. US20030118997A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephan
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 142.US5.REG
; CURRENT APPLICATION NUMBER: US/09/978,418
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/311,305
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/314,734
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/318,204
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/326,470
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: JPatent
; SEQ ID NO 40
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-418-40

Query Match 49.0%; Score 254; DB 10; Length 479;
Best Local Similarity 100.0%; Pred. No. 5.7e-240;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVNNKPTVKSLPCDLC 60
Db 7 LASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVNNKPTVKSLPCDLC 66

```
; PRIOR APPLICATION NUMBER: 08/958,970
; PRIOR FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 523
; TYPE: PRP
; ORGANISM: homo sapiens
US-09-767-007A-2

Query Match      36.1%; Score 187; DB 9; Length 523;
Best Local Similarity 100.0%; Pred. No. 2.9e-174;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLGAALAGPVLGLKECTRGSAVWQNVKTSASDCGAVKHCLQTVWKNKPTVKSLPCDIC 60
Db 7 LASLGAALAGPVLGLKECTRGSAVWQNVKTSASDCGAVKHCLQTVWKNKPTVKSLPCDIC 66

Qy 61 KDVVTAAGDMLKDNATBEELVYLEKTCDWLPKPNMGSASCKEIVDSYLPVILDIKEMS 120
Db 67 KDVVTAAGDMLKDNATBEELVYLEKTCDWLPKPNMGSASCKEIVDSYLPVILDIKEMS 126

Qy 121 RPGEVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVVAPFMANIPLILYPQDGP 180
Db 127 RPGEVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVVAPFMANIPLILYPQDGP 186

Qy 181 SKPOPKD 187
Db 187 SKPOPKD 193

RESULT 10
US-09-767-007A-3
; Sequence 3, Application US/09767007A
; Patent No. US2002007275A1
; GENERAL INFORMATION:
; APPLICANT: John S. O'Brien
; APPLICANT: Yasuo Kishimoto
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
; TITLE OF INVENTION: SAPOSIN C AND NEUROTROPIC PEPTIDES DERIVED THEREFROM
; FILE REFERENCE: MYELOS.2DC1C1
; CURRENT APPLICATION NUMBER: US/09/767,007A
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 08/958,970
; PRIOR FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 80
; TYPE: PRP
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide
US-09-767-007A-3

Query Match      15.4%; Score 80; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDVTYGSILS 364
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSEECQEVVDVTYGSILS 60

Qy 365 ILLEEVSPELVCSMLHLCSG 384
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 11
US-09-753-126-3
; Sequence 3, Application US/09753126
; Patent No. US20020127219A1
; GENERAL INFORMATION:
; APPLICANT: OKKELS, JENS SIGURD
; APPLICANT: JENSEN, ANNE DAM
; APPLICANT: HALKIER, TORBEN
; APPLICANT: JENSEN, RIKKE BOLDING
; TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME
; TITLE OF INVENTION: ACTIVATORS
; FILE REFERENCE: 31-0006000US
; CURRENT APPLICATION NUMBER: US/10/330,697
; CURRENT FILING DATE: 2002-12-27
; PRIOR APPLICATION NUMBER: US/09/753,126
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PA 1999 01891
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 60/174,652
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PA 200 00865
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/210,984
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 60/211,124
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: PA 2000 01027
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/217,497
```


; ORGANISM: Homo sapiens
US-09-780-438A-1

Query Match 7.7%; Score 40; DB 10; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.4e-31;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLP 344
|||
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLP 40

Search completed: May 5, 2004, 13:43:38
Job time : 41.7582 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:42:09 ; Search time 16.405 Seconds
(without alignments)
3037.320 Million cell updates/sec

Title: US-09-743-684A-1_COPY_7_524

Perfect score: 518

Sequence: 1 LASLGAALAGPVLGLKCT.....NTETAQCNAVEHCKRHVMN 518

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	278	53.7	527	1 SAHUP	saposin precursor
2	22	4.2	79	2 A49475	cerebroside sulfat
3	22	4.2	554	1 A28716	saposin precursor
4	22	4.2	557	1 JH0604	saposin precursor
5	13	2.5	80	2 S21770	saposin-C - bovine
6	9	1.7	81	2 A32026	glucosylceramide b
7	8	1.5	177	2 T47133	hypothetical prote
8	8	1.5	186	1 HHFF23	heat shock protein
9	8	1.5	186	2 B20647	heat shock protein
10	8	1.5	201	2 S36451	fimbrial protein f
11	8	1.5	201	2 S01929	fimbrial protein p
12	8	1.5	216	2 F64390	coenzyme F420 hydr
13	8	1.5	326	2 B90841	probable ABC trans
14	8	1.5	326	2 A85699	probable ABC trans
15	8	1.5	354	2 A23614	probable undecapre
16	8	1.5	435	2 D69992	ABC transporter (m
17	8	1.5	655	2 D85988	hypothetical prote
18	8	1.5	655	2 B65116	hypothetical 73.6
19	8	1.5	655	2 A91143	hypothetical prote
20	8	1.5	813	1 A49123	fibroblast growth
21	8	1.4	62	2 S77874	probable preprotei
22	7	1.4	94	2 A89986	hypothetical prote
23	7	1.4	105	2 S46149	ribosomal protein
24	7	1.4	112	2 A37203	lens fiber membran
25	7	1.4	125	2 B83857	chorismate mutase
26	7	1.4	148	2 JC2111	helix-loop-helix p
27	7	1.4	154	2 S76225	hypothetical prote
28	7	1.4	158	2 E84956	acetylactate synth
29	7	1.4	161	2 H84289	transcription regu

RESULT 1

SAHUP

Saposin precursor [validated] - human

N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component ein (SAP); sphingolipid activator protein A2; sulfate sulfatase activator protein N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D C;Species: Homo sapiens (man)

C;Date: 30-Jun-1992 #sequence revision 17-Nov-1995 #text change 08-Dec-2000

C;Accession: JX0061; A57368; A42003; B42003; C42003; D42003; A30367; S34740; S36140; S360226; I37265; I37264

R;Nakano, T.; Sandhoff, K.; Stuenkel, J.; Christomanou, H.; Suzuki, K.

J. Biochem. 105, 152-154, 1989

A;Title: Structure of full-length cDNA coding for sulfate activator, a Co-beta-glucosidase

A;Reference number: JX0061; MUID:89255151; PMID:2498298

A;Accession: JX0061

A;Molecule type: mRNA

A;Residues: 1-527 <NAK>

A;Cross-references: GB:D00422; NID:G220063; PIDN:BAA00321.1; PID:G220064

A;Note: alternative splice form 1

A;Accession: A57368

A;Molecule type: mRNA

A;Residues: 1-259,263-527 <NA2>

A;Cross-references: GB:J03015; GB:J03086; NID:G337755; PIDN:AAB59494.1; PID:G337756

A;Note: alternative splice form 2

R;Rorman, E.G.; Scheinker, V.; Grabowski, G.A.

Genomics 13, 312-318, 1992

A;Title: Structure and evolution of the human prosaposin chromosomal gene.

A;Reference number: A42003; MUID:92307663; PMID:1612590

A;Accession: A42003

A;Molecule type: DNA

A;Residues: 50-140 <ROR>

A;Cross-references: GB:M86181

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107236)

A;Accession: B42003

A;Molecule type: DNA

A;Residues: 185-259;263-276 <RO2>

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107237)

A;Accession: C42003

A;Molecule type: DNA

A;Residues: 305-393 <RO3>

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107238); sequence inc

A;Accession: D42003

A;Molecule type: DNA

A;Residues: 399-487 <RO4>

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107239); sequence inc

R;Rorman, E.G.; Grabowski, G.A.

Genomics 5, 486-492, 1989

A;Title: Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four spin

A;Reference number: A30367; MUID:90129043; PMID:2515150

A;Accession: A30367

A;Molecule type: mRNA

A;Residues: 1-259,263-527 <RO5>

A;Cross-references: GB:J03077; NID:G183230; PIDN:AAA52560.1; PID:G183231
 A;Note: alternative splice form 2
 R;Hiraawa, M.; O'Brien, J.S.; Kishimoto, Y.; Galdricka, M.; Fluharty, A.L.; Ginns, E.I.; Arch. Biochem. Biophys. 304, 110-116, 1993
 A;Title: Isolation, characterization, and proteolysis of human prosaposin, the precursor
 A;Reference number: S34740; MUID:93311991; PMID:8323276
 A;Accession: S34740
 A;Molecule type: protein
 A;Residues: 17-24,165-172;180-189;301-305 <HIR>
 R;Tyynelae, J.; Palmer, D.N.; Baumann, M.; Haltia, M. FEBS Lett. 330, 8-12, 1993
 A;Title: Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis.
 A;Reference number: S36140; MUID:93380576; PMID:8370464
 A;Accession: S36140
 A;Molecule type: protein
 A;Residues: 'XX',62,'X',64-65,'X',67-79,'X',81-84 <TY>
 A;Note: saposin A
 A;Accession: S36141
 A;Molecule type: protein
 A;Residues: 'XXX',413-414,'X',416-428,'X',430-434 <TY2>
 A;Note: saposin D
 R;Holtzman, H.; Sandhoff, K.; Kwon, H.Y.; Harzer, K.; Nakano, T.; Suzuki, K. J. Biol. Chem. 266, 7556-7560, 1991
 A;Title: Sulfatide activator protein. Alternative splicing that generates three mRNAs and
 A;Reference number: S36988; MUID:91210267; PMID:2019586
 A;Accession: S36988
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-240,'S',242-527 <HOL>
 A;Cross-references: EMBL:M60255; NID:G337759; PIDN:AAA36594.1; PID:G337760
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
 A;Note: cerebroside sulfate activator protein mutant MU-9; corresponds to alternative splicing
 A;Accession: S36989
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-240,'S',242-259,263-527 <HO2>
 A;Cross-references: EMBL:M60257; NID:G337764; PIDN:AAA36595.1; PID:G337765
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
 A;Note: cerebroside sulfate activator protein mutant MU-0; corresponds to alternative splicing
 A;Accession: S36990
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-240,'S',242-259,261-527 <HO3>
 A;Cross-references: EMBL:M60258; NID:G337766; PIDN:AAA36596.1; PID:G337767
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
 A;Note: cerebroside sulfate activator protein mutant MU-6; corresponds to alternative splicing
 R;Kondoh, K.; Hinenio, T.; Sano, A.; Kakimoto, Y. Biochem. Biophys. Res. Commun. 181, 286-292, 1991
 A;Title: Isolation and characterization of prosaposin from human milk.
 A;Reference number: PS0330; MUID:92068206; PMID:1958198
 A;Accession: PS0330
 A;Molecule type: protein
 A;Residues: 17-24,'X',26 <KON>
 A;Experimental source: milk
 R;Kretz, K.A.; Carson, G.S.; Morimoto, S.; Kishimoto, Y.; Fluharty, A.L.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 87, 2541-2544, 1990
 A;Title: Characterization of a mutation in a family with saposin B deficiency: a glycosyl
 A;Reference number: A35985; MUID:90207231; PMID:2320574
 A;Accession: A35985
 A;Molecule type: mRNA
 A;Residues: 213-221 <KR5>
 A;Cross-references: GB:M32221
 A;Accession: B35985
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-259,263-527 <KR2>
 A;Cross-references: GB:M32221; NID:G337761; PIDN:AAA60303.1; PID:G337762
 A;Experimental source: lymphoblast
 A;Accession: C35985
 A;Molecule type: mRNA
 A;Residues: 213-216,'I',218-221 <KR3>
 A;Note: sequence from patients with activator-deficient metachromatic leukodystrophy; th
 R;Fuerst, W.; Schubert, J.; Machleidt, W.; Meyer, H.E.; Sandhoff, K.

Eur. J. Biochem. 192, 709-714, 1990
 A;Title: The complete amino-acid sequences of human ganglioside GM2 activator protein and
 A;Reference number: S13195; MUID:9106165; PMID:2209618
 A;Accession: S13196
 A;Molecule type: protein
 A;Residues: 195-259,263-277 <FUE>
 R;Morimoto, S.; Martin, B.M.; Yamamoto, Y.; Kretz, K.A.; O'Brien, J.S.; Kishimoto, Y. Proc. Natl. Acad. Sci. U.S.A. 86, 3389-3393, 1989
 A;Title: Saposin A: second cerebroside activator protein.
 A;Reference number: A32784; MUID:89240739; PMID:2717620
 A;Accession: A32784
 A;Molecule type: protein
 A;Residues: 60-84;86-107;109-119;125-134 <MOR>
 R;O'Brien, J.S.; Kretz, K.A.; Dewji, N.; Wenger, D.A.; Esch, F.; Fluharty, A.L. Science 241, 1098-1101, 1988
 A;Title: Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic
 A;Reference number: A41240; MUID:88321660; PMID:2842863
 A;Accession: A41240
 A;Molecule type: mRNA
 A;Residues: 'GSSR',18-259,263-299,'D',301-302,'D',304-527 <OAB>
 A;Cross-references: GB:J03086
 R;Dewji, N.N.; Wenger, D.A.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 84, 8552-8556, 1987
 A;Title: Nucleotide sequence of cloned cDNA for human sphingolipid activator protein 1 p
 A;Reference number: S02289; MUID:88068647; PMID:2825202
 A;Accession: S02289
 A;Status: significant sequence differences
 A;Molecule type: mRNA
 A;Cross-references: EMBL:J03015
 A;Note: this sequence corrected by A41240
 R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 369, 1361-1365, 1988
 A;Title: Complete amino-acid sequence of the naturally occurring A(2) activator protein f
 A;Reference number: S02028; MUID:89207118; PMID:3242555
 A;Accession: S02028
 A;Molecule type: protein
 A;Residues: 195-259,263-276 <KLB>
 R;Fuerst, W.; Machleidt, W.; Sandhoff, K. Biol. Chem. Hoppe-Seyler 369, 317-328, 1988
 A;Title: The precursor of sulfatide activator protein is processed to three different pr
 A;Reference number: S00813; MUID:89000190; PMID:3048308
 A;Accession: S00813
 A;Molecule type: protein
 A;Residues: 410-487 <PU2>
 R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 368, 1571-1578, 1987
 A;Title: Complete amino-acid sequence and carbohydrate content of the naturally occurring
 A;Reference number: S00226; MUID:88163077; PMID:3442600
 A;Accession: S00226
 A;Molecule type: protein
 A;Residues: 314-333 <KL2>
 R;Vaccaro, A.M.; Salvioli, R.; Barca, A.; Tatti, M.; Ciaffoni, F.; Maras, B.; Siciliano, J. Biol. Chem. 270, 9953-9960, 1995
 A;Title: Structural analysis of saposin C and B. Complete localization of disulfide bridge
 A;Reference number: A57297; MUID:95247790; PMID:7730378
 A;Contents: annotation; disulfide bonds; glycosylation
 R;Holtzman, H.; Sandhoff, K.; Fuerst, W.; Kwon, H.Y.; Schnabel, D.; Suzuki, K. FEBS Lett. 280, 267-270, 1991
 A;Title: The organization of the gene for the human cerebroside sulfate activator protein
 A;Reference number: I37264; MUID:91192146; PMID:2013321
 A;Accession: I37265
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 59-125 <RES>
 A;Cross-references: EMBL:X57107; NID:G30234; PIDN:CAA40391.1; PID:G30235
 A;Accession: I37264
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 307-516 <RE2>
 A;Cross-references: EMBL:X57108; NID:G30232; PIDN:CAA40392.1; PID:G1565257
 A;Note: sequence revised relative to PID:G30233 (corrected coding region)
 C;Genetics:

Db 47 CKNVISQYSEIAIQMMHMQPK 68

RESULT 3

A28716

saposin precursor - rat

N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfata N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D

C;Species: Rattus norvegicus (Norway rat)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C;Accession: A28716

R;Collard, M.W.; Sylvester, S.R.; Tsuruta, J.K.; Griswold, M.D.

Biochemistry 27, 4557-4564, 1988

A;Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat s

A;Reference number: A28716; MUID:89000647; PMID:3048385

A;Accession: A28716

A;Molecule type: mRNA

A;Residues: 1-554 <COL>

A;Cross-references: GB:M1936; NID:G206904; PIDN:AAA42136.1; PID:G206905

A;Note: parts of this sequence, including the amino end of the mature protein, were dete

C;Function:

A;Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc

A;Pathway: sphingolipid catabolism

A;Note: saposin A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy

A;Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsulfa

A;Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiesterase

C;Superfamily: saposin; saposin repeat homology

C;Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph

F;1-16/Domain: signal sequence #status predicted <SIG>

F;17-554/Product: prosaposin #status predicted <PRO>

F;55-148/Domain: saposin repeat homology <SAP1>

F;60-143/Product: saposin A #status predicted <SAPA>

F;189-280/Domain: saposin repeat homology <SAP2>

F;194-273/Product: saposin B #status predicted <SAB1>

F;306-397/Domain: saposin repeat homology <SAP3>

F;310-389/Product: saposin C #status predicted <SAPC>

F;431-522/Domain: saposin repeat homology <SAP4>

F;437-514/Product: saposin D #status predicted <SAPD>

F;63-138, 66-132, 94-106, 439-512, 442-506, 470-481/Disulfide bonds: #status predicted

F;80-214, 331, 456/Binding site: carbohydrate (Asn)

F;197-270, 200-264, 229-240, 314-387, 317-381, 345-356/Disulfide bonds: #status predicted

Query Match 4.2%; Score 22; DB 1; Length 554;

Best Local Similarity 100.0%; Pred. No. 3.2e-14;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 425 EILAALEKGCSEFLPPYKQCD 446

. |||||

Db 461 EILAALEKGCSEFLPPYKQCD 482

RESULT 4

JH0604

saposin precursor - mouse

N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfata N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D

C;Species: Mus musculus (house mouse)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C;Accession: JH0604

R;Tsuda, M.; Sakiyama, T.; Endo, H.; Kitagawa, T.

Biochem. Biophys. Res. Commun. 184, 1266-1272, 1992

A;Title: The primary structure of mouse saposin.

A;Reference number: JH0604; MUID:92272718; PMID:1590788

A;Accession: JH0604

A;Molecule type: mRNA

A;Residues: 1-557 <TSU>

A;Cross-references: GB:S36200; NID:G249386; PIDN:AAB22175.1; PID:G249387

A;Experimental source: liver

C;Function:

A;Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc

A;Pathway: sphingolipid catabolism

A>Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosyl
A>Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsulfa
C:Superfamily: saposin; saposin; saposin repeat homology
F:1-16/Domain: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph
F:17-55/Product: signal sequence #status predicted <SIG>
F:55-148/Domain: prosaposin #status predicted <PRO>
F:60-143/Domain: saposin A #status predicted <SAP>
F:189-283/Domain: saposin B #status predicted <SAP>
F:194-276/Product: saposin repeat homology <SAP>
F:309-400/Domain: saposin repeat homology <SAP>
F:333-392/Product: saposin C #status predicted <SAP>
F:434-525/Domain: saposin repeat homology <SAP>
F:440-517/Product: saposin D #status predicted <SAPD>
F:63-138,66-132,94-106,197-273,200-267,229-240,317-390,320-384,348-359,442-515,445-509,4
F:80,214,334,379,459/Binding site: carbonylate (Asn) (covalent) #status predicted

Query Match 4.2%; Score 22; DB 1; Length 557;
Best Local Similarity 100.0%; Pred. No. 3.2e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 425 EILAALEKGCSELPDPYKQCD 446
|||||
DB 464 EILAALEKGCSELPDPYKQCD 485
|||||

RESULT 5
S21770
saposin-C - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S21770
R:Sano, A.; Mizuno, T.; Kondoh, K.; Hineno, T.; Ueno, S.; Kakimoto, Y.; Morita, N.
Biochim. Biophys. Acta 1120, 75-80, 1992
A:Title: Saposin-C from bovine spleen: complete amino acid sequence and relation between
A:Reference number: S21770; MUID:92207994; PMID:1554743
A:Accession: S21770
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-80 <SAP>
C:Superfamily: saposin; saposin repeat homology

Query Match 2.5%; Score 13; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 SPELVCSMLHCS 383
|||||
DB 67 SPELVCSMLHCS 79
|||||

RESULT 6
A32026
glucosylceramide beta-glucosidase activator protein SAP-2 - guinea pig
C:Species: Cavia porcellus (guinea pig)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 17-May-1996
C:Accession: A32026
R:Sano, A.; Radin, N.S.; Johnson, L.L.; Tarr, G.E.
J. Biol. Chem. 263, 19597-19601, 1988
A:Title: The activator protein for glucosylceramide beta-glucosidase from guinea pig liv
A:Reference number: A32026; MUID:89066787; PMID:3198642
A:Accession: A32026
A:Molecule type: protein
A:Residues: 1-81 <SAP>
C:Superfamily: saposin; saposin repeat homology
F:1-81/Domain: saposin repeat homology <SAP>

Query Match 1.7%; Score 9; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 351 CQEVWDYTG 359
|||||

DB 47 CQEVWDYTG 55
|||||

RESULT 7
T47133
hypothetical protein DKFZp761P19121.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47133
R:Pouetka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24374
A:Accession: T47133
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-177 <AAA>
A:Cross-references: EMBL:AL162013
A:Experimental source: adult amygdala; clone DKFZp761P19121
C:Genetics:
A>Note: DKFZp761P19121.1

Query Match 1.5%; Score 8; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 425 EILAALEK 432
|||||
DB 146 EILAALEK 153
|||||

RESULT 8
HHPF23
heat shock protein 23 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 16-Feb-1997
C:Accession: A02919
R:Ingolia, T.D.; Craig, E.A.
Proc. Natl. Acad. Sci. U.S.A. 79, 2360-2364, 1982
A:Title: Four small Drosophila heat shock proteins are related to each other and to mamme
A:Reference number: A93909; MUID:82248004; PMID:6285380
A:Accession: A02919
A:Molecule type: DNA
A:Residues: 1-186 <ING>
A>Note: the authors translated the codon CCC for residue 25 as Leu, CGA for residue 32 as
A>Note: the codon usage table proposed by the authors and the translation of the protein
C:Comment: This small heat shock protein is related to alpha crystallin.
C:Genetics:
A:Gene: FlyBase:Hsp23
A:Cross-references: FlyBase:FBgn0001224
A:Map position: 3L (67B)
C:Superfamily: alpha-crystallin
C:Keywords: heat shock; stress-induced protein

Query Match 1.5%; Score 8; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 MANIPLLL 173
|||||
DB 1 MANIPLLL 8
|||||

RESULT 9
B20647
heat shock protein 23 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 09-Sep-1987 #sequence_revision 09-Sep-1987 #text_change 21-Jul-2000
C:Accession: B20647
R:Southgate, R.; Ayme, A.; Voellmy, R.
J. Mol. Biol. 165, 35-57, 1983
A:Title: Nucleotide sequence analysis of the Drosophila small heat shock gene cluster at
A:Reference number: A92897; MUID:83189140; PMID:6302284

A:Accession: B20647
 A:Molecule type: DNA
 A:Residues: 1-186 <SOU>
 A:Cross-references: GB:V00210; NID:g8099; PIDN:CAA23494.1; PID:g8100
 C:Genetics:
 A:Gene: FlyBase:Hsp23
 A:Cross-references: FlyBase:FBgn0001224
 C:Superfamily: alpha-crystallin

Query Match 1.5%; Score 8; DB 2; Length 186;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 166 MANIPLLL 173
 |||||
 Db 1 MANIPLLL 8

RESULT 10

S36451
 fimbrial protein fimX - Bordetella bronchiseptica
 C:Species: Bordetella bronchiseptica
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999
 C:Accession: S36451
 R:Savelkoul, P.; de Kerf, D.; de Groot, L.; Willems, R.; Mooi, F.; van der Zeijst, B.; C
 submitted to the EMBL Data Library, February 1993
 A:Description: Characterization of fimX, a gene encoding a fimbrial subunit protein of B
 A:Reference number: S36451
 A:Accession: S36451
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-201 <SAV>
 A:Cross-references: EMBL:X74118; NID:g396788; PIDN:CAA52215.1; PID:g396789
 C:Genetics:
 A:Gene: fimX
 C:Superfamily: type 1 fimbrial protein

Query Match 1.5%; Score 8; DB 2; Length 201;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LLGAALAG 11
 |||||
 Db 7 LLGAALAG 14

RESULT 11

S01929
 fimbrial protein precursor - Bordetella pertussis
 C:Species: Bordetella pertussis
 C:Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 26-Aug-1999
 C:Accession: S01929
 R:Pedroni, P.; Riboli, B.; de Ferra, F.; Grandi, G.; Toma, S.; Arico, B.; Rappuoli, R.
 Mol. Microbiol. 2, 539-543, 1988
 A:Title: Cloning of a novel pilin-like gene from Bordetella pertussis: homology to the f
 A:Reference number: S01929; MUID:89013896; PMID:2902506
 A:Accession: S01929
 A:Molecule type: DNA
 A:Residues: 1-201 <PDP>
 A:Cross-references: EMBL:Y00556; NID:g39750; PIDN:CAA68634.1; PID:g39751
 C:Genetics:
 A:Gene: fimX
 C:Superfamily: type 1 fimbrial protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-201/Product: fimbrial protein #status predicted <MAT>

Query Match 1.5%; Score 8; DB 2; Length 201;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LLGAALAG 11
 |||||
 Db 7 LLGAALAG 14

RESULT 12

F64390
 coenzyme F420 hydrogenase (EC 1.12.99.1) gamma chain - Methanococcus jannaschii
 N:Alternate names: coenzyme F420-reducing hydrogenase, gamma subunit
 C:Species: Methanococcus jannaschii
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 02-Aug-2002
 C:Accession: F64390
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
 ; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
 raon, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A:Reference number: A64300; MUID:96337999; PMID:8688087
 A:Accession: F64390

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-216 <BUL>
 A:Cross-references: GB:U67519; GB:L77117; NID:gi591436; PIDN:AAB98722.1; PID:gi591442; T
 C:Genetics:

A:Map position: REV659511-658861
 C:Superfamily: coenzyme F420 hydrogenase gamma chain; ferredoxin 2[4Fe-4S] homology
 C:Keywords: oxidoreductase
 F:162-211/Domain: ferredoxin 2[4Fe-4S] homology <FER3>

Query Match 1.5%; Score 8; DB 2; Length 216;
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 KEIVDSYL 108
 |||||
 Db 117 KEIVDSYL 124

RESULT 13

B90841
 probable ABC transporter permease protein ECs1698 [imported] - Escherichia coli (strain
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C:Accession: B90841
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
 A:Reference number: A99829; MUID:21156231; PMID:11258796

A:Accession: B90841
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-326 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA835121.1; PID:gl3361163; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: ECs1698

Query Match 1.5%; Score 8; DB 2; Length 326;
 Best Local Similarity 100.0%; Pred. No. 8.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LLGAALAG 11
 |||||
 Db 63 LLGAALAG 70

RESULT 14

A85699
 probable ABC transporter permease protein Z1965 [imported] - Escherichia coli (strain O1
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001
 C:Accession: A85699
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: A85699
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-326 <STO>
A:Cross-references: GB:AE005174; NID:G12514905; PIDN:AGS6053.1; GSPDB:GN00145; UMGF:Z19
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z1965

Query Match 1.5%; Score 8; DB 2; Length 326;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LLGAALAG 11
|||||
Db 63 LLGAALAG 70

RESULT 15
AF3614
probable undecaprenyl-phosphate alpha-N-acetylglucosaminyltransferase (EC 2.4.1.1-) [impo
C:Species: *Brucella melitensis*
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 03-Feb-2003
C:Accession: AF3614
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A:Reference number: AD3252; PMID:11756688
A:Accession: AF3614
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-354 <KUR>
A:Cross-references: GB:AE008918; PIDN:AAL54081.1; PID:gl7985038; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI10839
A:Map position: 11
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 1.5%; Score 8; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LLGAALAG 11
|||||
Db 213 LLGAALAG 220

Search completed: May 5, 2004, 13:48:28
Job time : 17.405 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:35:24 ; Search time 10.4395 Seconds

(without alignments)
2583.672 Million cell updates/sec

Title: US-09-743-684a-1_COPY_7_524

Perfect score: 518
Sequence: 1 LASLGAALAGPVGLKECT.....NTETAQCNVHCKRHVMN 518Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	518	100.0	524	1	SAP_HUMAN
2	518	100.0	524	1	SAP_HUMAN
3	518	100.0	524	1	SAP_HUMAN
4	518	100.0	524	1	SAP_HUMAN
5	518	100.0	524	1	SAP_HUMAN
6	518	100.0	524	1	SAP_HUMAN
7	518	100.0	524	1	SAP_HUMAN
8	518	100.0	524	1	SAP_HUMAN
9	518	100.0	524	1	SAP_HUMAN
10	518	100.0	524	1	SAP_HUMAN
11	518	100.0	524	1	SAP_HUMAN
12	518	100.0	524	1	SAP_HUMAN
13	518	100.0	524	1	SAP_HUMAN
14	518	100.0	524	1	SAP_HUMAN
15	518	100.0	524	1	SAP_HUMAN
16	518	100.0	524	1	SAP_HUMAN
17	518	100.0	524	1	SAP_HUMAN
18	518	100.0	524	1	SAP_HUMAN
19	518	100.0	524	1	SAP_HUMAN
20	518	100.0	524	1	SAP_HUMAN
21	518	100.0	524	1	SAP_HUMAN
22	518	100.0	524	1	SAP_HUMAN
23	518	100.0	524	1	SAP_HUMAN
24	518	100.0	524	1	SAP_HUMAN
25	518	100.0	524	1	SAP_HUMAN
26	518	100.0	524	1	SAP_HUMAN
27	518	100.0	524	1	SAP_HUMAN
28	518	100.0	524	1	SAP_HUMAN
29	518	100.0	524	1	SAP_HUMAN
30	518	100.0	524	1	SAP_HUMAN
31	518	100.0	524	1	SAP_HUMAN
32	518	100.0	524	1	SAP_HUMAN
33	518	100.0	524	1	SAP_HUMAN

34	7	1.4	298	1	XERD_PSEAE	O9hxg6 pseudomonas
35	7	1.4	298	1	XERD_PSEBK	O8hmvo pseudomonas
36	7	1.4	309	1	Y209_MYCPN	P75485 mycoplasma
37	7	1.4	313	1	MRAM_PSEAE	O9hv25 pseudomonas
38	7	1.4	344	1	PUR5_NEIMA	O9jua2 neisseria m
39	7	1.4	344	1	PUR5_NEIMB	O9j280 neisseria m
40	7	1.4	388	1	RGSK_HUMAN	O76081 homo sapien
41	7	1.4	396	1	CAPI_STAVU	P39855 streptoloco
42	7	1.4	396	1	TCR3_ECOLI	P02981 escherichia
43	7	1.4	400	1	TCR8_PASMO	P51564 pasteurella
44	7	1.4	401	1	NIFS_ENTAG	O52069 enterobacte
45	7	1.4	436	1	PKI_HUMAN	O15118 homo sapien

ALIGNMENTS

RESULT 1
ID SAP_HUMAN STANDARD. PRT. 524 AA.
AC P07602; P07292; P15793; P78538; P78541; P78546; P78547; P78558;
AC Q92739; Q92740; Q92741; Q92742;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Proctivator polypeptide precursor [Contains: Saposin A (Protein A);
DE Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside
DE sulfate activator) (CSact) (Dispersin) (Sulfatide/GM1 activator);
DE Saposin C (Co-beta-glucosidase) (AI activator) (Glucosylceramidase
DE activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D
DE (Protein C) (Component C)].
GN PSAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RC MEDLINE=90129043; PubMed=2515150;
RA Roman E.G., Grabowski G.A.;
RT "Molecular cloning of a human co-beta-glucosidase cDNA: evidence that
RT four sphingolipid hydrolase activator proteins are encoded by single
RT genes in humans and rats";
RL Genomics 5:486-492(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=89255151; PubMed=2498298;
RA Nakano T., Sandhoff K., Struemper J., Christomanou H., Suzuki K.;
RT "Structure of full-length cDNA coding for sulfatide activator, a
RT co-beta-glucosidase and two other homologous proteins: two alternate
RT forms of the sulfatide activator";
RL J. Biochem. 105:152-154(1989).
RN [3]
RP SEQUENCE FROM N.A. (ISOROM SAP-MU-0).
RC TISSUE=Brain, Eye, and Skin;
RC MEDLINE=22388257; PubMed=12477932;
RA Struenser R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinzi P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Faney J., Helton E., Kerteman M., Madan A., Rodríguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalske U., Smallos D.E.,

- RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (4)
RP SEQUENCE OF 59-125 AND 304-513 FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=91192146; PubMed=2013321;
RA Holtschmidt H., Sandhoff K., Furst W., Kwon H.Y., Schnabel D.,
RA Suzuki K.;
RT "The organization of the gene for the human cerebroside sulfate
RT activator protein.";
RL FEBS Lett. 280:267-270(1991).
RN (5)
RP SEQUENCE OF 164-524 FROM N.A.
RA MEDLINE=8806647; PubMed=2825202;
RA Dewi N.N., Wenger D.A., O'Brien J.S.;
RT "Nucleotide sequence of cloned cDNA for human sphingolipid activator
RT protein 1 precursor.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8652-8656(1987).
RN (6)
RP PARTIAL SEQUENCE OF 60-142.
RA MEDLINE=89240739; PubMed=2717620;
RA Morimoto S., Martin B.M., Yamamoto Y., Kretz K.A., O'Brien J.S.,
RA Kishimoto Y.;
RT "Saposin A: second cerebroside activator protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:3389-3393(1989).
RN (7)
RP SEQUENCE OF 195-263 FROM N.A.
RA MEDLINE=86130593; PubMed=2868718;
RA Dewi N.N., Wenger D.A., Fujibayashi S., Donoviel M., Esch F.,
RA Hill F., O'Brien J.S.;
RT "Molecular cloning of the sphingolipid activator protein-1 (SAP-1),
RT the sulfate sulfate activator.";
RL Biochem. Biophys. Res. Commun. 134:989-994(1986).
RN (8)
RP SEQUENCE OF 195-274.
RC TISSUE=Kidney;
RA MEDLINE=91006165; PubMed=2209618;
RA Furst W., Schuber J., Machleidt W., Meyer H.E., Sandhoff K.;
RT "The complete amino-acid sequences of the naturally occurring A2 activator
RT protein and cerebroside sulfate activator protein.";
RL Eur. J. Biochem. 192:709-714(1990).
RN (9)
RP SEQUENCE OF 195-274.
RA MEDLINE=89207118; PubMed=3242555;
RA Kleinschmidt T., Christomanou H., Braunitzer G.;
RT "Complete amino-acid sequence of the naturally occurring A2 activator
RT protein for enzymic sphingomyelin degradation: identity to the
RT sulfate activator protein (SAP-1).";
RL Biol. Chem. Hoppe-Seyler 369:1361-1365(1988).
RN (10)
RP SEQUENCE OF 311-390.
RA MEDLINE=88163077; PubMed=3442600;
RA Kleinschmidt T., Christomanou H., Braunitzer G.;
RT "Complete amino-acid sequence and carbohydrate content of the
RT naturally occurring glycosylceramide activator protein (A1 activator)
RT absent from a new human Gaucher disease variant.";
RL Biol. Chem. Hoppe-Seyler 368:1571-1578(1987).
RN (11)
RP SEQUENCE OF 407-484.
RA MEDLINE=89000190; PubMed=3048308;
RA Furst W., Machleidt W., Sandhoff K.;
RT "The precursor of the sulfate activator protein is processed to three
RT different proteins.";
RL Biol. Chem. Hoppe-Seyler 369:317-328(1988).
RN (12)
RP PARTIAL SEQUENCE OF 405-484.
RA MEDLINE=89025876; PubMed=2845979;
RA Morimoto S., Martin B.M., Kishimoto Y., O'Brien J.S.;
RT "Saposin D: a sphingomyelinase activator.";
RL Biochem. Biophys. Res. Commun. 156:403-410(1988).
RN (13)
RP SEQUENCE OF 17-26.
RC TISSUE=Milk;
RA MEDLINE=92068206; PubMed=1958198;
RA Kondoh K., Hino T., Sano A., Kakimoto Y.;
RT "Isolation and characterization of prosaposin from human milk.";
RL Biochem. Biophys. Res. Commun. 181:286-292(1991).
RN (14)
RP PARTIAL SEQUENCE (SAPOSIN B), AND STRUCTURE OF CARBOHYDRATES.
RC TISSUE=urine;
RA MEDLINE=20032116; PubMed=10562467;
RA Fluharty A.L., Lombardo C., Louis A., Stevens R.L., Whitelegge J.P.,
RA Waring A.V., To T., Fluharty C.B., Paul K.F.;
RT "Preparation of the cerebroside sulfate activator (CSAct or saposin B)
RT from human urine.";
RL Mol. Genet. Metab. 68:391-403(1999).
RN (15)
RP STRUCTURE OF CARBOHYDRATE ON ASN-215.
RA MEDLINE=2110404; PubMed=11180632;
RA Paul K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,
RA Stevens R.L., Fluharty C.B., Fluharty A.L.;
RT "Structure of the asparagine-linked sugar chains of porcine kidney and
RT human urine cerebroside sulfate activator protein.";
RL J. Mass Spectrom. 35:1416-1424(2000).
RN (16)
RP SAPOSIN D DISULFIDE BONDS.
RA MEDLINE=99337688; PubMed=10406958;
RA Tatti M., Salvioli R., Ciaffoni F., Pucci P., Andolfo A.,
RA Amoresano A., Vaccaro A.M.;
RT "Structural and membrane-binding properties of saposin D.";
RL Eur. J. Biochem. 263:486-494(1999).
RN (17)
RP SAPOSIN B DISULFIDE BONDS.
RA MEDLINE=22398398; PubMed=12510003;
RA Ann V.E., Paul K.F., Whitelegge J.P., Higginson J., Fluharty A.L.,
RA Prive G.G.;
RT "Expression, purification, crystallization, and preliminary X-ray
RT analysis of recombinant human saposin B.";
RL Protein Expr. Purif. 27:186-193(2003).
RN (18)
RP MASS SPECTROMETRY.
RC TISSUE=Urine;
RA MEDLINE=99441404; PubMed=10510427;
RA Paul K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,
RA Kruchinsky A.N., Standing K.G., Waring A.V., Stevens R.L.,
RA Fluharty C.B., Fluharty A.L.;
RT "Cerebroside sulfate activator protein (Saposin B): chromatographic
RT and electrospray mass spectrometric properties.";
RL J. Mass Spectrom. 34:1040-1054(1999).
RN (19)
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 195-273, AND MUTAGENESIS OF
RP ILE-240.
RA MEDLINE=22406333; PubMed=12518053;
RA Ann V.E., Paul K.F., Whitelegge J.P., Fluharty A.L., Prive G.G.;
RT "Crystal structure of saposin B reveals a dimeric shell for lipid
RT binding.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:38-43(2003).
RN (20)
RP REVIEW ON MLD VARIANTS.
RA MEDLINE=95170731; PubMed=7866401;
RA Gieselmann V., Zlotogora J., Harris A., Wenger D.A., Morris C.P.;
RT "Molecular genetics of metachromatic leukodystrophy.";
RL Hum. Mutat. 4:233-242(1994).
RN (21)
RP VARIANT MLD ILE-217.
RA MEDLINE=90147748; PubMed=2302219;
RA Rafi M.A., Zhang X.-L., Degala G., Wenger D.A.;
RT "Detection of a point mutation in sphingolipid activator protein-1
RT leukodystrophy.";
RL Biochem. Biophys. Res. Commun. 166:1017-1023(1990).
RN (22)
RP SEQUENCE FROM N.A. AND VARIANT MLD ILE-217.
RA MEDLINE=90207231; PubMed=2320574;

Query Match 100.0%; Score 518; DB 1; Length 524;
 Best Local Similarity 100.0%; Pred. No. 0; Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLIGALAGVGLKECTRGSAVWCVKTAASDCGAKHCLQTVWKKPVPKSLPCDIC 60
 DB 7 LASLIGALAGVGLKECTRGSAVWCVKTAASDCGAKHCLQTVWKKPVPKSLPCDIC 66
 QY 61 KDVTAAAGDMLKDNATTEELIYLEKTCDMLEPKPKMSASCKEIVDSYLPVLIIDIKGEMS 120
 DB 67 KDVTAAAGDMLKDNATTEELIYLEKTCDMLEPKPKMSASCKEIVDSYLPVLIIDIKGEMS 126
 QY 121 RRGVCSALNLCESLOKHAELNHOKLESNKIPLEMDTEVAPAPMANIPLLYPODGR 180
 DB 127 RRGVCSALNLCESLOKHAELNHOKLESNKIPLEMDTEVAPAPMANIPLLYPODGR 186
 QY 181 SKPQKNDGVQCQDCIQWVTDIQTAVRTNSTFVQALVEHVEKCDRLGPGMADICKNYIS 240
 DB 187 SKPQKNDGVQCQDCIQWVTDIQTAVRTNSTFVQALVEHVEKCDRLGPGMADICKNYIS 246
 QY 241 QYSEIATQMMHMQPEKICALVGFCDVEKEMQTLVPAKVASKNVIVPALEIVEPIKHE 300
 DB 247 QYSEIATQMMHMQPEKICALVGFCDVEKEMQTLVPAKVASKNVIVPALEIVEPIKHE 306
 QY 301 VPAKSDVYCEVCEPELVKVTYKLIIDNNKTEKEILDAFDKWCSTLPSLSBECQEVVDYGS 360
 DB 307 VPAKSDVYCEVCEPELVKVTYKLIIDNNKTEKEILDAFDKWCSTLPSLSBECQEVVDYGS 366
 QY 361 SILSILIEVSEPELVCSMLHLCSTGRLPALTVHTVQPKDGCCEVCKLVGLDNLNLEKN 420
 DB 367 SILSILIEVSEPELVCSMLHLCSTGRLPALTVHTVQPKDGCCEVCKLVGLDNLNLEKN 426
 QY 421 STKQEIILALEKGSFLDPYQKCDQVAYEYEPVLIIELVEMDPSFVCLIKIGACPSAH 480
 DB 427 STKQEIILALEKGSFLDPYQKCDQVAYEYEPVLIIELVEMDPSFVCLIKIGACPSAH 486
 QY 481 KPLGTEKCIWGSYWCNTEETAACQNAVEHKKRRHVN 518
 DB 487 KPLGTEKCIWGSYWCNTEETAACQNAVEHKKRRHVN 524

RESULT 2
 SAP_BOVIN STANDARD; PRT; 525 AA.
 ID SAP_BOVIN
 AC P26779; Q9NRG4; Rel. 23, Created
 DT 01-AUG-1992 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Proactivator polypeptide precursor [Contains: Saposin A (Protein A);
 Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside
 sulfate activator) (CSAct) (Dispersin) (Sulfatide/GM1 activator);
 Saposin C (Co-beta-glucosidase) (A1 activator) (Glucosylceramidase
 activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D
 (Protein C) (Component C)].
 DE PSAP.
 GN PSAP.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANTS.
 RC TISSUE=Mammary gland;
 RA Azuma N., Yoshida K.;
 RT "RT-PCR cloning of bovine prosaposin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 312-391.
 RC TISSUE=Spleen;
 RX MEDLINE=92207994; PubMed=1554743;
 RA Sano A., Mizuno T., Kondoh K., Hieno S.-I., Kakimoto Y.,
 RA Morita N.;

RT "Saposin-C from bovine spleen; complete amino acid sequence and
 relation between the structure and its biological activity.";
 RL Biochim. Biophys. Acta 1120:75-80(1992).
 CC -I- FUNCTION: The lysosomal degradation of sphingolipids takes place
 CC by the sequential action of specific hydrolases. Some of these
 CC enzymes require specific low-molecular mass, non-enzymic proteins:
 CC the sphingolipids activator proteins (coproteins) (By similarity).
 CC -I- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
 CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
 CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
 CC Saposin C apparently acts by combining with the enzyme and acidic
 CC lipid to form an activated complex, rather than by solubilizing
 CC the substrate.
 CC -I- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
 CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1
 CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
 CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
 CC Saposin B forms a solubilizing complex with the substrates of the
 CC sphingolipid hydrolases (By similarity).
 CC -I- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
 CC activator (EC 3.1.4.12) (By similarity).
 CC -I- SUBUNIT: Saposin B is a homodimer (By similarity).
 CC -I- SUBCELLULAR LOCATION: Lysosomal.
 CC -I- PTM: This precursor is proteolytically processed to 4 small
 CC peptides, which are similar to each other and are sphingolipid
 CC hydrolase activator proteins (By similarity).
 CC -I- SIMILARITY: Contains 2 saposin A-type domains.
 CC -I- SIMILARITY: Contains 4 saposin B-type domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB036791; BA95677.1; -
 DR InterPro: IPR003119; SAPA.
 DR InterPro: IPR007856; SAPB_1.
 DR InterPro: IPR008138; SAPB_2.
 DR InterPro: IPR008140; SAPB_sub.
 DR InterPro: IPR008373; Saposin.
 DR InterPro: IPR008139; SaposinB.
 DR Pfam: PF02199; SAPA; 2.
 DR Pfam: PF05184; SAPB_1; 4.
 DR Pfam: PF03489; SAPB_2; 4.
 DR PRINTS: PR01797; SAPOSIN.
 DR ProDom: PD001732; SAPB_sub; 3.
 DR SMART: SMO0162; SAPA; 2.
 DR SMART: SMO0118; SAPB; 4.
 KW Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat.
 FT SIGNAL 1 16
 FT PROPEP 17 58
 FT CHAIN 60 142
 FT PROPEP 144 195
 FT CHAIN 196 275
 FT PROPEP 277 310
 FT CHAIN 312 392
 FT PROPEP 393 404
 FT CHAIN 406 487
 FT PROPEP 489 525
 FT CHAIN 525 54
 FT DOMAIN 54 142
 FT DOMAIN 142 194
 FT DOMAIN 194 276
 FT DOMAIN 276 312
 FT DOMAIN 312 393
 FT DOMAIN 393 406
 FT DOMAIN 406 487
 FT DOMAIN 487 525
 FT DISULFD 63 138
 FT DISULFD 66 132
 FT DISULFD 94 106
 FT DISULFD 199 272
 FT DISULFD 202 266

SAPOSIN A.
 SAPOSIN B.
 SAPOSIN C.
 SAPOSIN D.
 SAPOSIN-LIKE TYPE A 1.
 SAPOSIN-LIKE TYPE B 1.
 SAPOSIN-LIKE TYPE B 2.
 SAPOSIN-LIKE TYPE B 3.
 SAPOSIN-LIKE TYPE B 4.
 SAPOSIN-LIKE TYPE A 2.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.

FT DISULFID 231 242 BY SIMILARITY.
 FT DISULFID 316 389 BY SIMILARITY.
 FT DISULFID 319 383 BY SIMILARITY.
 FT DISULFID 347 358 BY SIMILARITY.
 FT DISULFID 410 483 BY SIMILARITY.
 FT DISULFID 413 477 BY SIMILARITY.
 FT DISULFID 441 452 BY SIMILARITY.
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 427 427 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 127 127 H -> R.
 FT VARIANT 260 263 MODK -> IIRIR.
 FT CONFLICT 317 317 E -> Q (IN REF. 2).
 FT CONFLICT 367 367 R -> S (IN REF. 1).
 SQ SEQUENCE 525 AA; 58120 MW; 293AFCDFB9C4FA99 CRC64;

Query Match 5.8%; Score 30; DB 1; Length 525;
 Best Local Similarity 100.0%; Pred. No. 8e-22;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 PGEVCSALNCESLQKHLAEINQKOLESN 151
 DB 128 PGEVCSALNCESLQKHLAEINQKOLESN 157

RESULT 3
 ID SAP PIG STANDARD; PRT; 80 AA.
 AC P8105;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Saposin B (Cerebroside sulfate activator) (CS-ACT) (Non-specific
 DE activator) (Sphingolipid activator protein 1) (SAP-1).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 NCBI_TaxID=9823;
 (1)
 RP SEQUENCE OF 1-79.
 RC TISSUE=Kidney;
 RX MEDLINE=93229506; PubMed=8471613;
 RA Stevens R.L., Pauli K.F., Conklin K.A., Green B.N., Fluharty A.L.;
 RT "Porcine cerebroside sulfate activator: further structural
 RT characterization and disulfide identification.";
 RL Biochemistry 32:4051-4059(1993).
 (2)
 RP SEQUENCE OF 1-64.
 RC TISSUE=Kidney;
 RX MEDLINE=92222651; PubMed=1562358;
 RA Fluharty A.L., Katona Z., Meek W.E., Frei K., Fowler A.V.;
 RT "The cerebroside sulfate activator from pig kidney: purification and
 RT molecular structure.";
 RL Biochem. Med. Metab. Biol. 47:66-85(1992).
 (3)
 RP STRUCTURE OF CARBOHYDRATE ON ASN-21.
 RX MEDLINE=21110404; PubMed=11180632;
 RA Pauli K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,
 RA Stevens R.L., Fluharty C.B., Fluharty A.L.;
 RT "Structure of the asparagine-linked sugar chains of porcine kidney and
 RT human urine cerebroside sulfate activator protein.";
 RL J. Mass Spectrom. 35:1416-1424(2000).
 (4)
 RP MASS SPECTROMETRY.
 RC TISSUE=Kidney;
 RX MEDLINE=94441404; PubMed=10510427;
 RA Pauli K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,
 RA Krutchnick A.N., Standing K.G., Waring A.J., Stevens R.L.,
 RA Fluharty C.B., Fluharty A.L.;
 RT "Cerebroside sulfate activator protein (Saposin B): chromatographic
 RT and electrospray mass spectrometric properties.";

RL J. Mass Spectrom. 34:1040-1054(1999).
 CC -1- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
 CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1
 CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
 CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
 CC Saposin B forms a solubilizing complex with the substrates of the
 CC sphingolipid hydrolases.
 CC -1- SUBUNIT: Saposin B is a homodimer (By similarity).
 CC -1- PTM: The one residue extended Saposin B-Val is only found in a
 CC minority of the chains.
 CC -1- SIMILARITY: Contains 1 saposin B-type domain.
 DR GLYCOSULEDB; P81405; -
 DR InterPro; IPR007856; SAPB_1.
 DR InterPro; IPR008138; SAPB_2.
 DR InterPro; IPR008373; Saposin.
 DR InterPro; IPR008139; SaposinB.
 DR Pfam; PF05184; SAPB_1; 1.
 DR Pfam; PF03489; SAPB_2; 1.
 DR PRINTS; PRO1797; SAPOSIN.
 DR SMART; SM00118; SAPB; 1.
 DR Glycoprotein; Sphingolipid metabolism.
 FT CHAIN 1 79
 FT CHAIN 1 80
 FT DOMAIN 1 80
 FT DISULFID 4 77
 FT DISULFID 7 71
 FT DISULFID 36 47
 FT CARBOHYD 21 21
 SQ SEQUENCE 80 AA; 8949 MW; EFTBA249B3E789C CRC64;
 N-LINKED (GLCNAC. . .) (COMPLEX).
 /FTID=CAR 000177.

Query Match 4.2%; Score 22; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 1.6e-14;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 CKNYISQSEIATQMMHMQPK 256
 DB 47 CKNYISQSEIATQMMHMQPK 68

RESULT 4
 ID SAP RAT STANDARD; PRT; 554 AA.
 AC P10560; O62841; O64190;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
 GN PSAP OR SGPI.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 (1)
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Sertoli cells;
 RX MEDLINE=8900647; PubMed=3048385;
 RA Collard M.W., Sylvester S.R., Tsauru J.K., Griswold M.D.;
 RT "Biosynthesis and molecular cloning of sulfated glycoprotein 1
 RT secreted by rat Sertoli cells: sequence similarity with the
 RT 70-kilodalton precursor to sulfatide/GM1 activator.";
 RL Biochemistry 27:4557-4564(1988).
 (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=96128541; PubMed=8573994;
 RA Morales C.R., El-Alfy M., Zhao Q., Igodora S.A.;
 RT "Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in
 RT Sertoli cells.";
 RL Histol. Histopathol. 10:1023-1034(1995).
 (3)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testicle;

```

RX MEDLINE=961752245; PubMed=8601692;
RA Morales C.R., El-Alfy M., Zhao Q., Igdonra S.A.;
RT "Expression and tissue distribution of rat sulfated glycoprotein-1
RT (prosaposin).";
CC J. Histochem. Cytochem. 44:327-337(1996).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: Contains 2 saposin A-type domains.
CC -1- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M19936; AAA42136.1; -.
CC EMBL; S81353; AAB36042.2; -.
CC EMBL; S81373; AAB36233.2; -.
CC PIR; A28716; A28716.
CC InterPro; IPR003119; SAPA.
CC InterPro; IPR007856; SAPB_1.
CC InterPro; IPR008138; SAPB_2.
CC InterPro; IPR008140; SAPB_sub.
CC InterPro; IPR008373; Saposin.
CC InterPro; IPR008139; SaposinB.
CC Pfam; PF02199; SAPA; 2.
CC Pfam; PF05184; SAPB_1; 3.
CC Pfam; PF03489; SAPB_2; 4.
CC PRINTS; PR01797; SAPOSIN.
CC ProDom; PD001732; SAPB_sub; 3.
CC SMART; SM00162; SAPA; 2.
CC SMART; SM00118; SAPB; 4.
CC Signal; Glycoprotein; Repeat.
CC -----
FT CHAIN 1 554 SULFATED GLYCOPROTEIN 1.
FT DOMAIN 21 54 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 61 138 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 193 274 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 310 391 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 435 516 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 521 554 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 521 554 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 197 270 BY SIMILARITY.
FT DISULFID 200 264 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 314 387 BY SIMILARITY.
FT DISULFID 317 381 BY SIMILARITY.
FT DISULFID 345 356 BY SIMILARITY.
FT DISULFID 439 512 BY SIMILARITY.
FT DISULFID 442 506 BY SIMILARITY.
FT DISULFID 470 481 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 115 115 P -> L (IN REF. 2 AND 3).
FT CONFLICT 299 299 D -> V (IN REF. 2).
FT CONFLICT 462 462 I -> R (IN REF. 3).
FT CONFLICT 527 527 W -> R (IN REF. 3).
FT CONFLICT 536 536 S -> M (IN REF. 3).
SQ SEQUENCE 554 AA; 61123 MW; DFE3F3A3A0520C6B CRC64;

```

```

RESULT 5
SAP_MOUSE STANDARD; PRT; 557 AA.
ID SAP_MOUSE 061207; 060861; 064006; 064219;
AC 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN PSAP OR SGP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBT_Txidp=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92272718; PubMed=1590788;
RA Tsuda M., Sakiyama T., Endo H., Kitagawa T.;
RT "The primary structure of mouse saposin.";
RL Biochem. Biophys. Res. Commun. 184:1266-1272(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94272317; PubMed=8003952;
RA Sprecher-Levy H., Orr-Urtreger A., Lona P., Horowitz M.;
RT "Murine prosaposin: expression in the reproductive system of a gene
RL implicated in human genetic disease.";
RN Cell. Mol. Biol. 40:233-233(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96084310; PubMed=8565332;
RA Cao Q.P., Crain M.R.;
RT "Expression of SGP-1 mRNA in preimplantation mouse embryos.";
RL Dev. Genet. 17:263-271(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/c; TISSUE=Liver;
RA Zhao Q.Q., Hay N.N., Morales C.R.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: Contains 2 saposin A-type domains.
CC -1- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S36200; AAB22175.1; -.
CC EMBL; S71616; AAB31059.1; -.
CC EMBL; U27399; AAB92567.1; -.
CC EMBL; U57999; AAB02595.1; -.
CC PIR; JH0604; JH0604.
CC MGD; MGI:97783; Psap.
CC InterPro; IPR003119; SAPA.
CC InterPro; IPR007856; SAPB_1.
CC InterPro; IPR008138; SAPB_2.
CC InterPro; IPR008140; SAPB_sub.
CC InterPro; IPR008373; Saposin.
CC Pfam; PF02199; SAPA; 2.
CC Pfam; PF05184; SAPB_1; 4.
CC Pfam; PF03489; SAPB_2; 4.
CC PRINTS; PR01797; SAPOSIN.
CC ProDom; PD001732; SAPB_sub; 3.
CC SMART; SM00162; SAPA; 2.
CC SMART; SM00118; SAPB; 4.
CC Signal; Glycoprotein; Repeat.
FT SIGNAL 1 16 BY SIMILARITY.

```

```

Oy 425 ETLAALKGCSPFLPDYPYQKCD 446
Db 461 ETLAALKGCSPFLPDYPYQKCD 482

```

```

Query Match 4.2%; Score 22; DB 1; Length 554;
Best Local Similarity 100.0%; Pred. No. 8.9e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

FT CHAIN 17 557 SULFATED GLYCOPROTEIN 1.
FT DOMAIN 21 54 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 59 142 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 193 277 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 313 519 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 438 557 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 524 557 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 197 273 BY SIMILARITY.
FT DISULFID 200 267 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 317 390 BY SIMILARITY.
FT DISULFID 320 384 BY SIMILARITY.
FT DISULFID 348 359 BY SIMILARITY.
FT DISULFID 442 515 BY SIMILARITY.
FT DISULFID 445 509 BY SIMILARITY.
FT DISULFID 473 484 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 83 83 Q -> E (IN REF. 2).
FT CONFLICT 158 158 I -> V (IN REF. 3).
FT CONFLICT 160 160 MISSING (IN REF. 3).
FT CONFLICT 171 172 MS -> SA (IN REF. 3).
FT CONFLICT 244 244 V -> L (IN REF. 2).
FT CONFLICT 254 254 M -> I (IN REF. 3).
FT CONFLICT 255 255 L -> W (IN REF. 2).
FT CONFLICT 260 262 MISSING (IN REF. 3).
FT CONFLICT 307 322 N -> D (IN REF. 3).
FT CONFLICT 322 322 F -> L (IN REF. 2).
FT CONFLICT 345 350 AL -> GV (IN REF. 1).
FT CONFLICT 367 367 G -> D (IN REF. 3).
FT CONFLICT 370 370 L -> Q (IN REF. 2).
FT CONFLICT 373 373 I -> D (IN REF. 3).
FT CONFLICT 391 391 A -> T (IN REF. 3).
FT CONFLICT 393 393 R -> L (IN REF. 3).
FT CONFLICT 406 406 A -> R (IN REF. 2 AND 3).
FT CONFLICT 430 430 P -> R (IN REF. 2).
FT CONFLICT 445 445 C -> F (IN REF. 4).
FT CONFLICT 448 448 L -> P (IN REF. 4).
SQ SEQUENCE 557 AA; 61422 MW; 134593E204993555 CRC64;

Query Match 4.2%; Score 22; DB 1; Length 557;
Best Local Similarity 100.0%; Pred. No. 9e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 425 EILALEKGSFLPPYQKCD 446
Db 464 EILALEKGSFLPPYQKCD 485

RESULT 6
SAP_CHICK STANDARD; PRT; 518 AA.
ID SAP_CHICK
AC 013035;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Proactivator polypeptide precursor [Contains: Saposin A; Saposin B;
DE Saposin C; Saposin D].
GN PSAP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 194-203.
RC TISSUE=Brain, and Liver;
RX MEDLINE=98129745; PubMed=9461526;

```

```
FT DISULFID 64 139 BY SIMILARITY.
FT DISULFID 67 133 BY SIMILARITY.
FT DISULFID 95 107 BY SIMILARITY.
FT DISULFID 197 273 BY SIMILARITY.
FT DISULFID 200 267 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 311 384 BY SIMILARITY.
FT DISULFID 314 384 BY SIMILARITY.
FT DISULFID 342 353 BY SIMILARITY.
FT DISULFID 403 476 BY SIMILARITY.
FT DISULFID 406 470 BY SIMILARITY.
FT DISULFID 434 445 BY SIMILARITY.
FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 94 94 R -> T (IN REF. 2).
FT CONFLICT 486 486 E -> D (IN REF. 2).
SQ SEQUENCE 518 AA; 57601 MW; B803000E891C3963 CRC64;
```

Query Match
Best Local Similarity 100.0%; Score 12; DB 1; Length 518;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 CKEIVDSYLPVI 111
Db 107 CKEIVDSYLPVI 118

RESULT 7

SAP_CAVPO STANDARD; PRT; 81 AA.

AC P20097;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Saposin C (CO-beta-glucosidase) (Glucosylceramidase activator)

DE (Sphingolipid activator protein 2) (SAP-2).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystriognath; Caviidae; Cavia.

NCBI_TaxID=10141;

RN [1]

RP SEQUENCE.

RC TISSUE=Liver;

RX MEDLINE=89066787; PubMed=3198642;

RA Sano A., Radin N.S., Johnson L.L., Tarr G.E.;

RT "The activator protein for glucosylceramide beta-glucosidase from guinea pig liver. Improved isolation method and complete amino acid sequence".

RU J. Biol. Chem. 263:19597-19601(1988).

CC -I- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).

CC Saposin C apparently acts by combining with the enzyme and acidic lipid to form an activated complex, rather than by solubilizing the substrate.

CC -I- SIMILARITY: Contains 1 saposin B-type domain.

CC PIR: A32026; A32026.

DR HSSP; P42210; IQDM.

DR InterPro; IPR007856; SAPB_1.

DR InterPro; IPR008138; SAPB_2.

DR InterPro; IPR008140; SAPB_sub.

DR InterPro; IPR008373; Saposin.

DR Pfam; PF05184; SAPB_1; 1.

DR Pfam; PF03489; SAPB_2; 1.

DR PRINTS; PR01797; SAPOSIN.

DR ProDom; PD001732; SAPB_sub; 1.

DR SMART; SM00116; SAPB_1.

KM Glycoprotein; Sphingolipid metabolism.

KM SAPOSIN-LIKE TYPE B.

FT DOMAIN 1 81 N-LINKED (GLCNAC. . .).

FT CARBOHYD 22 22

```
FT DISULFID 5 78 BY SIMILARITY.
FT DISULFID 8 72 BY SIMILARITY.
FT DISULFID 36 47 BY SIMILARITY.
SQ SEQUENCE 81 AA; 8852 MW; E564CE1F0A292596 CRC64;
```

Query Match
Best Local Similarity 100.0%; Score 9; DB 1; Length 81;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 351 COEVDITVG 359
Db 47 COEVDITVG 55

RESULT 8

HS23_DROME STANDARD; PRT; 186 AA.

AC P02516; Q9V5X5;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Heat shock protein 23.

GN HSP23 OR CG4463.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=83189140; PubMed=6302284;

RA Southgate R., Ayme A., Voellmy R.;

RT "Nucleotide sequence analysis of the Drosophila small heat shock gene cluster at locus 67B.";

RL J. Mol. Biol. 165:35-57(1983).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=82248004; PubMed=6285380;

RA Ingolia T.D., Craig E.A.;

RT "Four small Drosophila heat shock proteins are related to each other and to mammalian alpha-crystallin.";

RL Proc. Natl. Acad. Sci. U.S.A. 79:2360-2364(1982).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Suton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brannon R.C., Rogers J.-H.C., Blazey R.G., Chapple M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mkios G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkovec D., Botchan M.R., Bouck J., Brockstein P., Brodtier P.,

RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Doudon K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris D.M., Nelson D.L.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K., Nixon K., Nusskern D.R., Paclet J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reihert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skapski M.P., Smith T.,
 RA Splier E., Spirding A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtskas R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 [4]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Berkeley; TISSUE=Embryo;
 RX MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champagne M.,
 RA George R.A., Guarini H., Krommiller B., Pacleb J.W., Park S., Wan K.H.,
 RA Rubin G.M., Celinker S.E.,
 RT "A *Drosophila* full-length cDNA resource,"
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 CC -1- SIMILARITY: Belongs to the small heat shock protein (HSP20)
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J01100; AAA28637.1; -
 DR EMBL: V00210; CA223494.1; -
 DR EMBL: X03889; CA27525.1; -
 DR EMBL: AE003552; AAF50286.1; -
 DR EMBL: AY061081; AAL28629.1; -
 DR PIR: B20647; B20647.
 DR FLYBASE: FBgn0001224; Hsp23.
 DR InterPro: IPR001436; Crystallin_alpha.
 DR InterPro: IPR002068; Hsp20.
 DR InterPro: IPR008978; Hsp20_chap.
 DR Pfam: PF00011; HSP20_1.
 DR PRINTS: PRO0299; ACRYSTALLIN.
 DR PROSITE: PS01031; HSP20_1.
 DR Heat shock; Multigene family.
 CC CONFLICT 31 33 ORN -> RRI (IN REF. 2).
 FT CONFLICT 81 81 K -> E (IN REF. 2).
 FT CONFLICT 88 88 K -> G (IN REF. 2).
 FT CONFLICT 95 95 L -> V (IN REF. 2).
 FT CONFLICT 173 173 N -> S (IN REF. 2).
 FT CONFLICT 182 182 N -> G (IN REF. 2).
 SQ SEQUENCE 186 AA; 20629 MW; 3CEEDDA1719B5A0 CRC64;
 Query Match
 Best Local Similarity 1.5%; Score 8; DB 1; Length 186;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 166 MANIPULL 173
 Db 1 MANIPULL 8
 RESULT 9
 FMFX BORPE STANDARD; PRT; 201 AA.
 AC P09808;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Fimbrial protein fimx precursor (Pilin).
 GN FIMX OR BP2674.
 OS *Bordetella pertussis*.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Alcaligenaceae; *Bordetella*.
 OX NCBI_TaxID=520;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SAI.
 RX MEDLINE=89013966; PubMed=2902506;
 RA Pedroni P., Riboli B., de Ferra F., Grandi G., Toma S., Arico B.,
 RA Rappuoli R.,
 RT "Cloning of a novel pilin-like gene from *Bordetella pertussis*:
 RT Mol. Microbiol. 2:539-543(1988).
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Tokama I / ATCC BAA-589 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebatina M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jags K.,
 RA Leatherbarrow S., Moul S., Norberg-Zakh H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin I., Whitehead S., Barrett B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
 RT *Bordetella parapertussis* and *Bordetella bronchiseptica*,"
 RL Nat. Genet. 35:32-40(2003).
 CC -1- SUBCELLULAR LOCATION: Pilin structure on the cell surface.
 CC -1- DISEASE: *Bordetella pertussis* is the causative agent of whooping
 CC cough. An essential step in the disease process is the attachment
 CC of the bacteria to the ciliated epithelium of the respiratory
 CC tract, enabling the organism to resist normal host clearance
 CC mechanisms. It is unclear which bacterial cell surface component
 CC are responsible for adherence but the fimbriae of *B. pertussis* are
 CC prime candidates for being involved in this process.
 CC -1- SIMILARITY: BELONGS TO THE FIMB/PAPA FAMILY OF FIMBRIA PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y00556; CA68634.1; -
 DR EMBL: A25464; CA01770.1; -
 DR EMBL: BX640419; CA642950.1; -
 DR PIR: S01929; S01929.
 DR InterPro: IPR008966; Adhes_bact.
 DR InterPro: IPR000259; Fimbrial.
 DR Pfam: PF00419; Fimbrial_1.
 KW Fimbria; Signal; Complete proteome.
 FT SIGNAL 1 21
 FT CHAIN 22 201 FIMBRIAL PROTEIN FIMX.
 FT DISULFID 37 79 PROBABLE.
 FT CONFLICT 165 166 AX -> GO (IN REF. 1).
 SQ SEQUENCE 201 AA; 21462 MW; F47670CC05B03F6 CRC64;
 Query Match
 Best Local Similarity 1.5%; Score 8; DB 1; Length 201;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 4 LIGALANG 11
 Db 7 LIGALANG 14
 RESULT 10
 FMFX METUA STANDARD; PRT; 216 AA.
 AC Q58136;
 ID Y726 METUA
 DT 05-11-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Fimbrial protein fimx precursor (Pilin).
 GN FIMX OR BP2674.
 OS *Bordetella pertussis*.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein MU0726.
 GN MU0726.
 OS Methanococcus jannaschii.
 OC Archaeae; Euryarchaeota; Methanococci; Methanococcales;
 CC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
 RA Utermbeck T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RT Science 273:1058-1073(1996).
 RL -1- SIMILARITY: SOME, TO COENZYME F420 HYDROGENASE GAMMA SUBUNIT.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U67519; AAB98722.1; -.
 DR PIR: F64390; F64390.
 DR TIGR: MJ0726; -.
 DR InterPro: IPR001450; 4FE4S_ferredoxin.
 DR InterPro: IPR006137; Oxidored_g6.
 DR Pfam: PF00037; fer4; 1.
 DR Pfam: PF01058; oxidored_g6; 1.
 DR PROSITE: PS00198; 4FE4S_FERREDOXIN; 1.
 KW Hypothetical protein; Oxidoreductase; Iron-sulfur; 4Fe-4S;
 KW Electron transport; Complete proteome.
 FT METRL 169 169 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
 FT METRL 172 172 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
 FT METRL 175 175 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
 FT METRL 179 179 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
 FT METRL 197 197 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
 FT METRL 200 200 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
 FT METRL 203 203 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
 FT METRL 207 207 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
 SO SEQUENCE 216 AA; 23789 MW; 2FDP067698982F6 CRC64.
 Query Match 1.5%; Score 8; DB 1; Length 216;
 Best Local Similarity 100.0%; Pred. No. 4.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 101 KEIYDSTL 108
 |||||
 DB 117 KEIYDSTL 124
 RESULT 11
 BORA_MOUSE STANDARD; PRT; 349 AA.
 AC Q9JMS6; Q9JMS6; 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cdc42 effector protein 4 (Bander of Rho GTPase 4).
 GN CDC42EP4 OR BORG4 OR CEP4.
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE=ileal mucosa;
 RX MEDLINE=21036164; PubMed=1185749;
 RA Osada N., Kusuda J., Suzuki Y., Sugano S., Hashimoto K.;
 RT "Sequence analysis, gene expression, and chromosomal assignment of
 RT mouse BORG4 gene and its human orthologue.";
 RL J. Hum. Genet. 45:374-377 (2000).
 [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=22388257; PubMed=12477932;
 RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casarini T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshitsuki S., Carrinci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting M., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywanski M.I., Skalska U., Smalins D.E.,
 RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [3]
 RP SEQUENCE OF 1-143 FROM N.A., AND INTERACTION WITH TC10 AND CDC42.
 RX MEDLINE=99421943; PubMed=10490598;
 RA Joberty G., Perlungher R.R., Macara I.G.;
 RT "The Borge, a new family of Cdc42 and Tc10 GTPase-interacting
 RT proteins.";
 RL Mol. Cell. Biol. 19:6585-6597(1999).
 CC -1- FUNCTION: Probably involved in the organization of the actin
 CC cytoskeleton. May act downstream of CDC42 to induce actin
 CC filament assembly leading to cell shape changes. Induces
 CC pseudopodia formation, when overexpressed in fibroblasts.
 CC -1- SUBUNIT: Interacts with CDC42 and Tc10, in a GTP-dependent manner.
 CC -1- SIMILARITY: BELONGS TO THE BORG/CEP FAMILY.
 CC -1- SIMILARITY: Contains 1 CRIB domain.
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- SIMILARITY: BELONGS TO THE BORG/CEP FAMILY.
 CC -1- SIMILARITY: Contains 1 CRIB domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB035088; BAA95932.1; -.
 DR EMBL: BC003857; AA03857.1; -.
 DR EMBL: AF165114; AAD7822.1; -.
 DR MGD: WGI:1929760; Cdc42ep4.
 DR GO: GO:0005737; C:cytoplasm; IDA.
 DR GO: GO:0005515; F:protein binding; IPI.
 DR GO: GO:0007266; P:Rho protein signal transduction; IDA.
 DR InterPro: IPR000995; PAKbox/Rho-binding.
 DR Pfam: PF00786; PBD; 1.
 DR PROSITE: PS50108; CRIB; 1.
 KW Cytoskeleton.
 FT DOMAIN 27 41 CRIB.
 FT CONFLICT 143 143 K -> R (IN REF. 3).

SQ SEQUENCE 349 AA; 37869 MW; 7C4125A7083E1B CRC64;
 Query Match 1.5%; Score 8; DB 1; Length 349;
 Best Local Similarity 100.0%; Pred.No. 6.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 341 SKLPSKLS 348
 Db 131 SKLPSKLS 138
 RESULT 12
 BOR4_HUMAN STANDARD; PRT; 356 AA.
 ID BOR4_HUMAN
 AC Q9H3Q1; Q95828; Q96FT3;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cdc42 effector protein 4 (Binder of Rho GTPase 4).
 GN CDC42BP4 OR BOR4 OR CEP4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ileal mucosa;
 RX MEDLINE=21036164; PubMed=1185749;
 RA Oseada N., Kusuda J., Suzuki Y., Sugano S., Hashimoto K.,
 RT "Sequence analysis, gene expression, and chromosomal assignment of
 RT mouse Borg4 gene and its human orthologue."
 RL J. Hum. Genet. 45:374-377(2000).
 [2]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC TISSUE=Eye;
 RX MEDLINE=21125609; PubMed=11035016;
 RA Hirsch D.S., Pirone D.M., Burbelo P.D.,
 RT "A new family of cdc42 effector proteins, CEPs, function in fibroblast
 RT and epithelial cell shape changes."
 RL J. Biol. Chem. 276:875-883(2001).
 [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=2238257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci F., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Muzny D.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Shvachenko Y., Bouffard G.G.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Dickson M.C.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skala U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [1]
 RP FUNCTION: Probably involved in the organization of the actin
 CC cytoskeleton. May act downstream of CDC42 to induce actin
 CC filament assembly leading to cell shape changes. Induces
 CC pseudopodia formation, when overexpressed in fibroblasts.
 [2]
 RP SUBUNIT: Interacts with CDC42 and TC10, in a GTP-dependent manner
 CC (by similarity).
 [3]
 RP SUBCELLULAR LOCATION: Cytoplasmic; membrane-associated.
 CC TISSUE SPECIFICITY: Not detected in any of the adult tissues
 CC tested. May be expressed only in fetal or embryonic tissues.

CC -1- SIMILARITY: BELONGS TO THE BOR4/CEP FAMILY.
 CC -1- SIMILARITY: Contains 1 CRIB domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC
 DR EMBL; AB042237; BAB17272.1;
 DR EMBL; AF099664; AAD16299.1;
 DR EMBL; BC010451; AAH10451.1;
 DR Genbank; HGNC:17147; CDC42BP4.
 DR MIM; 605468;
 DR InterPro; IPR000095; PAKbox/Rho-binding.
 DR Pfam; PF00766; PBD, 1.
 DR PROSITE; PS50108; CRIB, 1.
 KM Cysteine-rich.
 FT DOMAIN 27 41 CRIB.
 FT DOMAIN 70 74 POLY-SER.
 FT CONFLICT 2 2 P -> I (IN REF. 3).
 FT CONFLICT 33 33 D -> T (IN REF. 2).
 FT CONFLICT 288 288 A -> T (IN REF. 2).
 FT CONFLICT 296 296 A -> T (IN REF. 2).
 FT CONFLICT 339 339 P -> PP (IN REF. 2).
 FT CONFLICT 355 355 MISSING (IN REF. 2).
 SQ SEQUENCE 356 AA; 37979 MW; 2CF677C60C6E1F5 CRC64;
 Query Match 1.5%; Score 8; DB 1; Length 356;
 Best Local Similarity 100.0%; Pred.No. 6.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 341 SKLPSKLS 348
 Db 133 SKLPSKLS 140
 RESULT 13
 MNTC_BACSU STANDARD; PRT; 435 AA.
 ID MNTC_BACSU
 AC O35024;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Manganese transport system membrane protein mntC.
 GN MNTC OR BSU30750.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98048467; PubMed=9387221;
 RA Laplante A., Galleron N., Sorokin A., Ehrlich S.D.,
 RT "Sequencing and functional annotation of the Bacillus subtilis genes
 RT in the 200 kb rimb-dnaB region."
 RL Microbiology 143:3431-3441(1997).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=9804033; PubMed=9384377;
 RA Kunze F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Baetoro M.G., Besieres P., Borchert S., Bron S.,
 RA Brouillet S., Bruchet C.V., Brans A., Braun M., Brignell S.C., Carter N.M.,
 RA Choi S.K., Codan U.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denton K.D., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 RA Eutian K.D., Frixton J., Fabret C., Ferrari E., Foulger D.,
 RA Frits C.Y., Fujita M., Fujita Y., Fuma S., Gallazzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaute A.,

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karmata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardino S., Lauder J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigone C.,
RA Medina N., Meliada R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portelle D., Porollik S., Prescott A.M.,
RA Priesen E., Puje P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolet C., Roche E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlon E., Schleich S., Schroeter R., Scofield F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Solido B.,
RA Sorokin A., Taccini E., Takagi T., Takahashi H., Takemaru K.,
RA Takouchi M., Tamkooshi A., Tanaka T., Terpeira P., Tononi A.,
RA Toesto V., Uchiyama S., Vandenbol M., Vannier P., Vasarotti A.,
RA Viati A., Wandt R., Wedler E., Wedler H., Wetzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zumbstein A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*,"
RL Nature 390:249-256 (1997).
[3]
RP POSSIBLE FUNCTION.
RX MEDLINE=2023631; PubMed=10760146;
RA Que O., Helmann J.D.;
RT "Manganese homeostasis in *Bacillus subtilis* is regulated by MntR, a
RT difunctional regulator related to the diphtheria toxin repressor
RT family of proteins."
RL Mol. Microbiol. 35:1454-1468 (2000).
CC -1- FUNCTION: This protein is probably a component of a manganese
CC peremease, a binding protein-dependent, ATP-driven transport
CC system.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: BELONGS TO THE ABC-3 SUBFAMILY OF INTEGRAL MEMBRANE
CC PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL, AF008220; AAC00231.1; -
DR EMBL; Z91119; CAB15053.1; -
DR PIR; D69992; D69992.
DR Subtilist; BG13853; mntC.
DR InterPro: IPR001626; ABC_transp3.
DR Pfam; PF00950; ABC-3; 1.
KW Transport; Ion transport; Transmembrane; Complete proteome.
FT TRANSMEM 17 37 POTENTIAL.
FT TRANSMEM 42 62 POTENTIAL.
FT TRANSMEM 68 88 POTENTIAL.
FT TRANSMEM 98 118 POTENTIAL.
FT TRANSMEM 143 163 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT TRANSMEM 228 248 POTENTIAL.
FT TRANSMEM 255 275 POTENTIAL.
SQ SEQUENCE 435 AA; 47944 MW; 68B6D590D90A73BC CRC64;
Query Match 1.5%; Score 8; DB 1; Length 435;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ID FA10 RABIT STANDARD; PRT; 490 AA.
AC 019045;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).
GN F10
OS *Oryctolagus cuniculus* (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97256311; PubMed=9101642;
RA Pendurthi U.R., Anderson K.D., James H.L.;
RT "Characterization of a full-length cDNA for rabbit factor X";
RL Thromb. Res. 85:503-514 (1997).
CC -1- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that
CC converts prothrombin to thrombin in the presence of factor Va,
CC calcium and phospholipid during blood clotting.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then
CC Arg-|-Ile bonds in prothrombin to form thrombin.
CC -1- SUBUNIT: The two chains are formed from a single-chain precursor
CC by the excision of two Arg residues and are held together by 1 or
CC more disulfide bonds.
CC -1- PTM: The vitamin K-dependent, enzymatic carboxylation of some
CC glutamate residues allows the modified protein to bind calcium (By
CC similarity).
CC -1- PTM: N- and O-glycosylated (By similarity).
CC -1- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY)
CC (BY SIMILARITY).
CC -1- MISCELLANEOUS: Calcium also binds, with stronger affinity to
CC another site, beyond the Gla domain.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 2 EGF-like domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL, AF003200; AAB62542.1; -
DR HSSP; P00742; IHCG.
DR MEROPS; S01.216; -
DR InterPro: IPR000152; Aex_hydroxyl_S.
DR InterPro: IPR009003; Cys_Ser_trypsin.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_II.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR002383; Gla_blood.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1A.
DR InterPro: IPR000294; Vitk_dep_Gla.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00094; Gla; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00010; EGFBILOD.
DR PRINTS; PR00001; GLABLOOD.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00069; Gla; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00010; AEX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.

```

DR PROSITE; PS50240; TRYPSIN DOM; 1.
DR PROSITE; PS00134; TRYPSIN HIS; 1.
DR PROSITE; PS00135; TRYPSIN SER; 1.
KW Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;
KW Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;
KW Signal; Zymogen; EGF-like domain; Repeat.
FT SIGNAL 1 20
FT PROPEP 21 40
FT CHAIN 41 180
FT CHAIN 184 490
FT PROPEP 184 490
FT CHAIN 233 490
FT CHAIN 233 490
FT DOMAIN 86 122
FT DOMAIN 125 165
FT DOMAIN 233 490
FT MOD_RES 46 46
FT MOD_RES 47 47
FT MOD_RES 47 47
FT MOD_RES 54 54
FT MOD_RES 54 54
FT MOD_RES 56 56
FT MOD_RES 59 59
FT MOD_RES 60 60
FT MOD_RES 65 65
FT MOD_RES 66 66
FT MOD_RES 69 69
FT MOD_RES 72 72
FT MOD_RES 75 75
FT MOD_RES 79 79
FT MOD_RES 103 103
FT ACT_SITE 274 274
FT ACT_SITE 320 320
FT ACT_SITE 417 417
FT DISULFID 90 101
FT DISULFID 95 110
FT DISULFID 112 121
FT DISULFID 129 140
FT DISULFID 136 149
FT DISULFID 151 164
FT DISULFID 172 340
FT DISULFID 239 244
FT DISULFID 259 275
FT DISULFID 388 402
FT DISULFID 413 441
FT CARBOHYD 61 61
FT CARBOHYD 187 187
FT CARBOHYD 205 205
SQ SEQUENCE 490 AA; 53965 MW; 3A39FA85F2A6D11 CRC64;

Query Match 1.5%; Score 8; DB 1; Length 490;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 LGALAG 11
DB 9 LGALAG 16

```

```

DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein yncp.
GN YHCP OR B3240.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
NCBI_TaxID=562;
(1)
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655.
RX MEDLINE=9742617; PubMed=9278503;
RA Blaxter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: SOME, TO B.CEPACIA FUSC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC -----
DR EMBL; U18997; AA56042.1; -.
DR EMBL; AB000403; AAC76272.1; -.
DR PIR; B65116; B65116.
DR Ecogen; EG12822; yncp.
DR InterPro; IPR006726; FUSC.
DR Pfam; PF04632; FUSC; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 13 33
FT TRANSMEM 38 58
FT TRANSMEM 69 89
FT TRANSMEM 93 113
FT TRANSMEM 121 141
FT TRANSMEM 152 172
FT TRANSMEM 370 390
FT TRANSMEM 407 427
FT TRANSMEM 431 451
FT TRANSMEM 459 479
FT TRANSMEM 482 502
SQ SEQUENCE 655 AA; 73591 MW; 6C559FC19C276656 CRC64;

Query Match 1.5%; Score 8; DB 1; Length 655;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 74 NATEEIL 81
DB 331 NATEEIL 338

```

Search completed: May 5, 2004, 13:46:13
Job time : 11.4395 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:41:29 ; Search time 37.7812 Seconds
(without alignments)
4325.921 Million cell updates/sec

Title: US-09-743-684a-1_COPY_7_524

Perfect score: 518
Sequence: 1 LASLGAALAGPVGLKECT.....NTETAAQCNAYEHCKRHWN 518

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_todent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacterioph:*
- 17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	4.2	554	11	Q8BFQ1
2	20	3.9	121	6	P79254
3	13	2.5	512	13	Q7SY70
4	12	2.3	449	11	Q8BJV5
5	12	2.3	525	11	Q8C1C1
6	11	2.1	402	11	Q8C1C1
7	9	1.7	520	13	Q8UVZ4
8	9	1.7	522	13	Q8D682
9	8	1.5	50	9	Q854G1
10	8	1.5	177	4	Q9NSM6
11	8	1.5	201	2	Q44889
12	8	1.5	201	16	Q7WGY5
13	8	1.5	294	13	Q98S10
14	8	1.5	297	16	Q7VPN8
15	8	1.5	300	2	Q9JUN6
16	8	1.5	300	16	Q8D492

17	8	1.5	326	16	Q8X584	Q8X584 escherichia
18	8	1.5	326	16	Q8XRP7	Q8XRP7 shigella fl
19	8	1.5	330	2	Q8X7G2	Q8X7G2 methylobact
20	8	1.5	337	12	Q8UTY7	Q8UTY7 lumpy skin
21	8	1.5	337	12	Q8UTN0	Q8UTN0 lumpy skin
22	8	1.5	337	12	Q91ML5	Q91ML5 lumpy skin
23	8	1.5	354	16	Q8YB05	Q8YB05 brucella me
24	8	1.5	373	2	Q937R1	Q937R1 brucella me
25	8	1.5	375	16	Q8FWM3	Q8FWM3 brucella su
26	8	1.5	390	5	Q862X0	Q862X0 plasmodium
27	8	1.5	405	16	Q8XSP7	Q8XSP7 raietonia s
28	8	1.5	456	2	Q93CS1	Q93CS1 shigella bo
29	8	1.5	655	16	Q8X9E6	Q8X9E6 escherichia
30	8	1.5	655	16	Q8FDS1	Q8FDS1 escherichia
31	8	1.5	655	16	Q83JF2	Q83JF2 shigella fl
32	8	1.5	663	2	Q841B5	Q841B5 neisseria m
33	8	1.5	775	5	Q81JP7	Q81JP7 plasmodium
34	8	1.5	1160	16	Q8BDZ4	Q8BDZ4 pseudomonas
35	7	1.4	13	4	Q75905	Q75905 homo sapien
36	7	1.4	55	16	Q7VBC3	Q7VBC3 prochloroxo
37	7	1.4	61	2	Q490E5	Q490E5 mycoplasma
38	7	1.4	68	9	Q856S0	Q856S0 mycobacteri
39	7	1.4	80	16	Q8RC04	Q8RC04 thermococci
40	7	1.4	83	2	Q9AED2	Q9AED2 lactobacilli
41	7	1.4	90	9	Q9ZXL7	Q9ZXL7 bacterioph
42	7	1.4	94	16	Q9S8S9	Q9S8S9 seaphylococ
43	7	1.4	101	17	Q8PMG5	Q8PMG5 methanococ
44	7	1.4	102	16	Q9YXK4	Q9YXK4 streptococc
45	7	1.4	102	16	Q8E2R0	Q8E2R0 streptococc

ALIGNMENTS

RESULT 1

ID	Q8BFQ1	PRELIMINARY:	PRT:	554 AA.
AC	Q8BFQ1			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Proseposin.			
GN	PSAP.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N. A.			
RC	STRAIN=NOB; TISSUE=Kidney, and Thymus;			
RX	MEDLINE=22354683; PubMed=12466851;			
RA	The FANTOM Consortium,			
RA	"Analysis of the mouse transcription Research Group Phase I & II Team;			
RT	60,770 full-length cDNAs."			
RL	Nature 420:563-573(2002).			
DR	EMBL; AK088369; BAC40308.1; -			
DR	EMBL; AK089998; BAC41035.1; -			
DR	MGD; MGI:97783; Psap.			
DR	GO; GO:0005764; C:lysosome; IEA.			
DR	GO; GO:0006655; P:sphingolipid metabolism; IEA.			
DR	InterPro; IPR003119; SAPA.			
DR	InterPro; IPR007856; SAPB_1.			
DR	InterPro; IPR008138; SAPB_2.			
DR	InterPro; IPR008140; SAPB_sub.			
DR	InterPro; IPR008373; Saposin.			
DR	InterPro; IPR008139; Saposinb.			
DR	Pfam; PF02199; SAPA_2.			
DR	Pfam; PF05184; SAPB_1; 4.			
DR	Pfam; PF03489; SAPB_2; 4.			
DR	PRINTS; PR01797; SAPOSIN.			
DR	ProDom; PD001732; SAPB_sub; 3.			
DR	SMART; SMO0162; SAPA_2.			

DR SMART; SM00118; SAPB; 4.
SQ SEQUENCE 554 AA; 61050 MW; FFS8DB79C7CC0C18 CRC64;

Query Match
Best Local Similarity 4.2%; Score 22; DB 11; Length 554;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 425 EILAALEKGCSEFLDPYOKOCD 446
|||||
DB 461 EILAALEKGCSEFLDPYOKOCD 482

RESULT 2

ID P79254 PRELIMINARY; PRT; 121 AA.

AC P79254;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DE Sulfated glycoprotein-1/SGP-1 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]

RP SEQUENCE FROM N.A.
RA MEDLINE=96249303; PubMed=8848570;
RA Spencer T.E., Graf G.H., Bazer F.W.;
RT "Sulfated glycoprotein-1 (SGP-1) expression in ovine endometrium during the oestrous cycle and early pregnancy."
RL Reprod. Fert. Dev. 7:1053-1060(1995).
DR EMBL: S82555; AAD1405.1; -
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR008138; SAPB 2.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF03489; SAPB_2; 1.
DR PRINTS; PR01797; SAPOSIN.
DR SMART; SM00118; SAPB; 1.
FT NON TER
FT SEQUENCE 121 AA; 13604 MW; 4F0F5A6EB83D0C9A CRC64;

Query Match
Best Local Similarity 3.9%; Score 20; DB 6; Length 121;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 219 HYKEBCDRLGPGMADICKNY 238
|||||
DB 21 HYKEBCDRLGPGMADICKNY 40

RESULT 3

ID Q7SY70 PRELIMINARY; PRT; 512 AA.

AC Q7SY70;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8955;
RN [1]

RP SEQUENCE FROM N.A.
RA TISSUE=Whole;
RA MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus

RT Initiative";
RT Dev. Dyn. 225:384-391 (2002).

RN [2]
RP SEQUENCE FROM N.A.
RA TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heist F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Prange C.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalka U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RA TISSUE=Whole;
RA Klein S., Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC054988; AAH54988.1; -
DR Hypothetical protein.
SQ SEQUENCE 512 AA; 57555 MW; 57CF5A5E2093FEB1 CRC64;

Query Match
Best Local Similarity 2.5%; Score 13; DB 13; Length 512;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 VNNKPTVKSLPCD 58
|||||
DB 52 VNNKPTVKSLPCD 64

RESULT 4

ID Q8BV5 PRELIMINARY; PRT; 449 AA.

AC Q8BV5;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical saposin A-type domain/saposin type B containing
DE protein.
GN 2310020A21RIK.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Eye;
RA MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).

DR EMBL; AK078699; BAC37363.1; -
DR GDB; MGI:1924193; 2310020A21RIK.
DR GO; GO:0005764; C:lysosome; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SAPB_1.

```
DR InterPro; IPR008138; SAPB_2.
DR InterPro; IPR008140; SAPB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SABA_1.
DR Pfam; PF05184; SAPB_1; 3.
DR Pfam; PF03489; SAPB_2; 3.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SAPB_sub; 2.
DR SMART; SM00162; SABA_1.
DR SMART; SM00118; SAPB; 3.
DR Hypothetical protein.
SQ SEQUENCE 449 AA; 48705 MW; 377CC42A475B292F CRC64;

Query Match
Best Local Similarity 2.3%; Score 12; DB 11; Length 449;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 VWNKPTVKSLLPC 57
Db 56 VWNKPTVKSLLPC 67

RESULT 5
08C1C1 PRELIMINARY; PRT; 525 AA.
ID 08C1C1
AC 08C1C1
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical saposin A-type domain/saposin type B containing
DE protein.
GN 2310020A21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK028455; BAC25961.1; -.
DR MGD; MGI:1924193; 2310020A21RIK.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006655; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; SABA.
DR InterPro; IPR007856; SAPB_1.
DR InterPro; IPR008138; SAPB_2.
DR InterPro; IPR008140; SAPB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SABA_1; 2.
DR Pfam; PF05184; SAPB_1; 4.
DR Pfam; PF03489; SAPB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SAPB_sub; 3.
DR SMART; SM00162; SABA_2.
DR SMART; SM00118; SAPB; 4.
DR Hypothetical protein.
SQ SEQUENCE 525 AA; 57350 MW; 374F6050CDC4D223 CRC64;

Query Match
Best Local Similarity 2.3%; Score 12; DB 11; Length 525;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 VWNKPTVKSLLPC 57
Db 56 VWNKPTVKSLLPC 67
```

```
RESULT 6
08C1N0 PRELIMINARY; PRT; 402 AA.
ID 08C1N0
AC 08C1N0
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical saposin A-type domain/saposin type B containing
DE protein.
GN 2310020A21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK009408; BAC25258.1; -.
DR MGD; MGI:1924193; 2310020A21RIK.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006655; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; SABA.
DR InterPro; IPR007856; SAPB_1.
DR InterPro; IPR008138; SAPB_2.
DR InterPro; IPR008140; SAPB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SABA_1.
DR Pfam; PF05184; SAPB_1; 3.
DR Pfam; PF03489; SAPB_2; 3.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SAPB_sub; 2.
DR SMART; SM00162; SABA_1.
DR SMART; SM00118; SAPB; 3.
DR Hypothetical protein.
SQ SEQUENCE 402 AA; 44420 MW; E90017CBF4017ED6 CRC64;

Query Match
Best Local Similarity 2.1%; Score 11; DB 11; Length 402;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 448 FVAEYEPVLI 458
Db 322 FVAEYEPVLI 332

RESULT 7
08UVZ4 PRELIMINARY; PRT; 520 AA.
ID 08UVZ4
AC 08UVZ4
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Lysosomal cofactor/neurotrophic factor prosaposin.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteiophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC Seo H.-C., Lie O., Fjose A., O'Brien J.S., Kishimoto Y.;
RT "Cloning, expression and promoter analysis of zebrafish prosaposin.";
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
EMBL; AF276996; AAL54381.1; -.
```

DR GO; GO:0005764; C:lysosome; IEA.
 DR GO; GO:0006655; P:sphingolipid metabolism; IEA.
 DR InterPro; IPR003119; SAPA.
 DR InterPro; IPR007856; SAPB_1.
 DR InterPro; IPR008138; SAPB_2.
 DR InterPro; IPR008140; SAPB_sub.
 DR InterPro; IPR008373; Saposin.
 DR InterPro; IPR008139; SaposinB.
 DR Pfam; PF02199; SAPA; 2.
 DR Pfam; PF05184; SAPB_1; 3.
 DR Pfam; PF03489; SAPB_2; 4.
 DR PRINTS; PR01797; SAPOSIN.
 DR PRODOM; PD001732; SAPB_sub; 3.
 DR SMART; SM00162; SAPA; 2.
 DR SMART; SM00118; SAPB; 4.
 SQ SEQUENCE 520 AA; 57431 MW; F9E620F84BA1CB5 CRC64;

Query Match 1.7%; Score 9; DB 13; Length 520;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 WCONVKTAS 34
 DB 30 WCONVKTAS 38

RESULT 8
 ID Q9DG82 PRELIMINARY; PRT; 522 AA.
 AC Q9DG82;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Prosaposin.
 GN PSAP.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_TaxID=7955;
 OX
 RN
 RP SEQUENCE FROM N.A.
 RA Altman N., Horowitz M.;
 RT "The zebrafish prosaposin cDNA."
 RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF108655; AAC32919.1; --
 DR ZFIN; ZDB-GENE-020108-1; psap.
 DR GO; GO:0005764; C:lysosome; IEA.
 DR GO; GO:0006655; P:sphingolipid metabolism; IEA.
 DR InterPro; IPR003119; SAPA.
 DR InterPro; IPR007856; SAPB_1.
 DR InterPro; IPR008138; SAPB_2.
 DR InterPro; IPR008140; SAPB_sub.
 DR InterPro; IPR008373; Saposin.
 DR InterPro; IPR008139; SaposinB.
 DR Pfam; PF02199; SAPA; 2.
 DR Pfam; PF05184; SAPB_1; 3.
 DR Pfam; PF03489; SAPB_2; 4.
 DR PRINTS; PR01797; SAPOSIN.
 DR PRODOM; PD001732; SAPB_sub; 3.
 DR SMART; SM00162; SAPA; 2.
 DR SMART; SM00118; SAPB; 4.
 SQ SEQUENCE 522 AA; 57671 MW; D3C15A305725C1CD CRC64;

Query Match 1.7%; Score 9; DB 13; Length 522;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 WCONVKTAS 34
 DB 30 WCONVKTAS 38

RESULT 9
 ID O854G1 PRELIMINARY; PRT; 50 AA.
 AC O854G1;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Gp106.
 GN 106.
 OS Mycobacteriophage Omega.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
 OC NCBI_TaxID=205879;
 OX
 RN
 RP SEQUENCE FROM N.A.
 RA MEDLINE=22592660; PubMed=12705866;
 RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
 RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Panunzio N.R.,
 RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
 RA Krizakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
 RA Hatfull G.F.;
 RT "Origins of highly mosaic mycobacteriophage genomes."
 RT Cell 113:171-182(2003).
 RL EMBL; AY129338; AAN12747.1;
 DR EMBL; AY129338; AAN12747.1;
 SQ SEQUENCE 50 AA; 5689 MW; 16993D01C4E247F4 CRC64;

Query Match 1.5%; Score 8; DB 9; Length 50;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 ILDIKGE 118
 DB 43 ILDIKGE 50

RESULT 10
 ID Q9NSM6 PRELIMINARY; PRT; 177 AA.
 AC Q9NSM6;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein (Fragment).
 GN DKFZ761P19121.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 OX
 RN
 RP SEQUENCE FROM N.A.
 RA TISSUE=Amalgam;
 RA Pousetka A., Melienreuther R., Mewes H.W., Weil B., Wiemann S.;
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL162013; CAB82365.1; --
 DR PIR; T47133; T47133.
 KW Hypothetical protein.
 FT NON_TER
 SQ SEQUENCE 177 AA; 20522 MW; EB717DE9993AF417 CRC64;

Query Match 1.5%; Score 8; DB 4; Length 177;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 425 ETLALEX 432
 DB 146 ETLALEX 153

RESULT 11
 ID Q44889 PRELIMINARY; PRT; 201 AA.
 AC Q44889;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Fimbrial subunit precursor.
 GN FIMX.
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NO. 401;
 RA Savelkoul P., De Keif D., de Groot L., Willems R., Mooi F.,
 RA van der Zeijst B., Gaastera W.;
 RT "Characterization of fimbria, a gene encoding a fimbrial subunit protein
 RT of Bordetella bronchiseptica."
 RL Submitted (FEB-1993) to the EMBL/Genbank/DBJ databases.
 DR EMBL: X74118; CA52215.1; -.
 DR PIR: S36451; S36451.
 DR GO: GO:0009289; C:fimbria; IEA.
 DR GO: GO:0007155; P:cell adhesion; IEA.
 DR InterPro: IPR008966; Adhes.bact.
 DR InterPro: IPR000259; Fimbrial.
 DR Pfam: PF00419; Fimbrial; 1.
 DR Signal.
 KW Signal.
 FT SIGNAL.
 SQ SEQUENCE 201 AA; 21462 MW; F47670CC05B03F6 CRC64;

Query Match 1.5%; Score 8; DB 2; Length 201;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LGALAG 11
 |||||
 Db 7 LGALAG 14

RESULT 12
 O7MGYS PRELIMINARY; PRT; 201 AA.
 ID O7MGYS.
 AC O7MGYS.
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Fimbrial protein.
 GN FIMX OR BB3426.
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RB50 / ATCC BAA-588;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parthill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Fellwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
 RA Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica."
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL: BX640447; CAE33918.1; -.
 KW Complete proteome.
 SQ SEQUENCE 201 AA; 21511 MW; CB4970CC05439082 CRC64;

Query Match 1.5%; Score 8; DB 16; Length 201;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LGALAG 11
 |||||
 Db 7 LGALAG 14

RESULT 13
 O98S10 PRELIMINARY; PRT; 294 AA.
 ID O98S10.
 AC O98S10.
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Hoxala protein.
 GN HOXALA.
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McClintock J.M.;
 RT "Consequences of Hox gene duplication in the vertebrates: an
 RT investigation of the zebrafish hox paralogue group 1 genes."
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 DR EMBL: AJ306431; CAC34566.1; -.
 DR HSSP; P14653; 1B72.
 DR ZFIN; ZDB-GENE-000823-5; hoxala.
 DR GO: GO:0005634; C:nucleus; IEA.
 DR GO: GO:0003700; F:transcription factor activity; IEA.
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IPR001356; Homeobox.
 DR InterPro: IPR000047; HTH_LambdaRepressor.
 DR Pfam: PF00046; homeobox; 1.
 DR PRINTS; PR00024; HOMEBOX.
 DR PRINTS; PR00031; HTHREPRESSR.
 DR ProDom; PD000010; Homeobox; 1.
 DR SMART; SM00389; HOX; 1.
 DR PROSITE; PS00027; HOMEBOX_1; 1.
 DR PROSITE; PS50071; HOMEBOX_2; 1.
 KW DNA-binding; Homeobox; Nuclear protein.
 SQ SEQUENCE 294 AA; 32179 MW; 8A7C2391BF410727 CRC64;

Query Match 1.5%; Score 8; DB 13; Length 294;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 343 LPKSLSEE 350
 |||||
 Db 254 LPKSLSEE 261

RESULT 14
 O7VPN8 PRELIMINARY; PRT; 297 AA.
 ID O7VPN8.
 AC O7VPN8.
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Integrase/recombinase Xerd.
 GN XERD OR HD0013.
 OS Haemophilus ducreyi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=730;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=35000HP / ATCC 700724;
 RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
 RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
 RT "The complete genome sequence of Haemophilus ducreyi."
 RL Submitted (JUN-2003) to the EMBL/Genbank/DBJ databases.

DR EMBL; AE017151; AAP95036.1; -
 KW Complete proteome.
 SQ SEQUENCE 297 AA; 34472 MW; 589A1B3984E73951 CRC64;

Query Match
 Best Local Similarity 1.5%; Score 8; DB 16; Length 297;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 343 LPKSLSEE 350
 |||||
 Db 108 LPKSLSEE 115

RESULT 15

Q93UN6 PRELIMINARY; PRT; 300 AA.
 ID Q93UN6
 AC Q93UN6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
 DE Hypothetical protein.
 OS Vibrio vulnificus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=672;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CM-1;
 RA Chang M.C., Chang S.Y., Chang C.C.;
 RT "Partial DNA sequence of the chromosome in Vibrio vulnificus."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DDO databases.
 CC -1 SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 DR EMBL; AF080431; AAK50454.1; -
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription; DNA-dependent; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR InterPro; IPR000847; HTH_LYR.
 DR InterPro; IPR005119; LysR_subst.
 DR Pfam; PF00126; HTH_1; 1.
 DR Pfam; PF03466; LysR_substrate; 1.
 DR KW Hypothetical protein; DNA-binding; Transcription;
 KW transcription regulation.
 SQ SEQUENCE 300 AA; 33710 MW; B7B508680F6D8ED CRC64;

Query Match
 Best Local Similarity 1.5%; Score 8; DB 2; Length 300;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LIGALAG 11
 |||||
 Db 223 LIGALAG 230

Search completed: May 5, 2004, 13:47:42
 Job time : 38.7812 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:15:48 ; Search time 72.3134 Seconds
(without alignments)
2023.963 Million cell updates/sec

Title: US-09-743-684A-1_COPY_7_524
Perfect score: 2759
Sequence: 1 LASILGALAGPVLGKRECT.....NTEFAAGCAVEHCKRHVYN 518

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2759	100.0	524	2	AAR70783 Prosaposin
2	2759	100.0	524	2	AAW85652 Human pro
3	2759	100.0	524	3	AAy58716 Human pro
4	2759	100.0	524	6	ABU79099 Lip-TAA b
5	2759	100.0	524	6	ABU05200 Human exp
6	2759	100.0	524	6	ABU05207 Human exp
7	2759	100.0	524	6	ABU05203 Human exp
8	2759	100.0	524	6	ABU07340 Human exp
9	2759	100.0	524	6	ABU05216 Human exp
10	2759	100.0	524	6	ABU05202 Human exp
11	2759	100.0	524	6	ABU05214 Human exp
12	2759	100.0	524	6	ABU05215 Human exp
13	2759	100.0	524	6	ABU05199 Human exp
14	2759	100.0	524	6	ABU05212 Human exp
15	2759	100.0	524	6	ABU05213 Human exp
16	2759	100.0	524	6	ABU05205 Human exp
17	2749	99.6	524	6	ABU05208 Human exp
18	2747.5	99.6	527	5	AAAB31915 Amino aci
19	2747.5	99.6	527	5	ABE68602 Human pan
20	2747.5	99.6	527	6	ABU79100 Lip-TAA b
21	2747.5	99.6	527	6	ABU05204 Human exp
22	2747.5	99.6	527	6	ABU05210 Human exp
23	2742.5	99.4	523	4	AAAB31916 Amino aci
24	2742.5	99.4	523	6	ABU05211 Human exp
25	2738	99.2	526	6	ABU05209 Human exp

ALIGNMENTS

RESULT 1									
ID	AAAR70783	standard;	protein;	524	AA.				
AC	XX								
XX	AAAR70783;								
DT	25-MAR-2003	(revised)							
DT	30-AUG-1995	(first entry)							
XX									
DE	Prosaposin.								
XX									
KW	Saposin-C; neuron; myelination; nervous system; neuroblastoma;								
KW	neurotrophic peptide; multiple sclerosis; leukoencephalitis;								
KW	adrenal leukodystrophy; prosaposin.								
OS	Human sapiens.								
PN	W09503821-A1.								
PD	09-FEB-1995.								
PF	28-JUL-1994;	94WO-US008453.							
PR	30-JUL-1993;	93US-00100247.							
PR	21-APR-1994;	94US-00232513.							
PA	(O'BRIEN / O'BRIEN J S.								
PI	O'brien JS, Kishimoto Y;								
DR	MPJ. 1995-082029/11.								
DR	N-PSDB; AAQ85355.								
PT	Stimulating neural cell out-growth and myelination - with pro:saposin,								
PT	saposin C or new neurotrophic peptide(s) from cytokine(s), for treating								
XX	nervous system diseases.								
PS	Disclosure: Page 30-32; 50pp; English.								
XX									
CC	The peptide given in AAR70773, corresponding to amino acids 8-29 of human								
CC	saposin-C (AAR70784), promotes neurite outgrowth in vitro. A consensus								
CC	sequence was determined by comparing the peptide with hematopoietic and								
CC	neurotrophic cytokines, and neurotrophic peptides (AAR70774-82) were								
CC	identified in the AB loop of human ciliary neurotrophic factor,								
CC	interleukins-6, -2, -3 and -gamma, erythropoietin and leukocyte								
CC	inhibitory factor, and in helix C of human interleukin-1-beta and								
CC	oncostatin-M. Prosaposin (AAR70783) and saposin-C also promoted nerve								
CC	cell myelination ex vivo. (Updated on 25-MAR-2003 to correct PN field.)								

CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to
 CC correct PI field.)

XX Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 2; Length 524;
 Best Local Similarity 100.0%; Pred. No. 7.2e-235;

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 LASLLGALLAGPVGLGKCTRGSAVWCQNVKTASDCGAVHCLQTVWKNKPTVKSLLPCDIC 60
DB 7 LASLLGALLAGPVGLGKCTRGSAVWCQNVKTASDCGAVHCLQTVWKNKPTVKSLLPCDIC 66
QY 61 KDVTAAAGDMLKDNATEEELIVLEKTCMDLPKPNMSASCKEIVDSYLPVILDIITIKGEMS 120
DB 67 KDVTAAAGDMLKDNATEEELIVLEKTCMDLPKPNMSASCKEIVDSYLPVILDIITIKGEMS 126
QY 121 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPEMANIPLLLYPODGPR 180
DB 127 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPEMANIPLLLYPODGPR 186
QY 181 SKPQKNDGVCDQCIQMTVDIQTAVRTNSTFVQALVEHKEECDRLGPGMADICKNYIS 240
DB 187 SKPQKNDGVCDQCIQMTVDIQTAVRTNSTFVQALVEHKEECDRLGPGMADICKNYIS 246
QY 241 QYSEIATQMMHMQPEKICALVGFCDVEKEMPQTLVPAKVASKNVVPALTEVEPIKXHE 300
DB 247 QYSEIATQMMHMQPEKICALVGFCDVEKEMPQTLVPAKVASKNVVPALTEVEPIKXHE 306
QY 301 VAKSDVYCEVEEFLVKEVTKILINNKTKEKILDAFDKXCSLPKSLSEECQEVVDITYGS 360
DB 307 VAKSDVYCEVEEFLVKEVTKILINNKTKEKILDAFDKXCSLPKSLSEECQEVVDITYGS 366
QY 361 SILSILIEEVSPELVCSMLHCSGTRLPALTVAHVQPKDGFCEVCKLVGYLDRNLEKN 420
DB 367 SILSILIEEVSPELVCSMLHCSGTRLPALTVAHVQPKDGFCEVCKLVGYLDRNLEKN 426
QY 421 STKQSIILALEKGSFLLPDPYQKCDQFAVEEPEVLIEIIVMDPSFVCLIKAGCSAH 480
DB 427 STKQSIILALEKGSFLLPDPYQKCDQFAVEEPEVLIEIIVMDPSFVCLIKAGCSAH 486
QY 481 KPLLSTEKICIMGPSYWCQNTERRAACNAVEHCKRHVN 518
DB 487 KPLLSTEKICIMGPSYWCQNTERRAACNAVEHCKRHVN 524

```

RESULT 2

AAW85652 ID AAW85652 standard; protein; 524 AA.

XX AAW85652;

DT 19-JUL-1999 (first entry)

DE Human prosaposin N-terminal peptide.

XX Prosaposin; saposin; prosaptilides; prosaposin receptor agonists; PRA;
 XX peripheral nervous system; central nervous system; PNS; CNS; Akt; Bcl-2;
 XX therapy; treatment; apoptosis; caspase; tumor necrosis factor; TNF;
 XX cytokine; interleukin gamma; IFN; inflammation; rheumatoid arthritis;
 XX Crohn's disease; irritable bowel syndrome; ashma; cardiac infarction;
 XX congestive heart failure; multiple sclerosis;
 XX acute disseminated inflammatory leukoencephalitis;
 XX progressive multifocal leukoencephalitis; Alzheimer's disease;
 XX Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;
 XX ischemic heart disease; Guillain-Barre disease; alopecia; AIDS dementia;
 XX cerebral malaria; HIV; neuropathy;
 XX inflammatory neurodegenerative disease; toxin-induced liver disease.

OS Homo sapiens.

XX MO9912559-A1.

PD 18-MAR-1999.

XX 09-SEP-1998; 98WO-US019216.

XX 09-SEP-1997; 97US-0058352P.

XX 04-JUN-1998; 98US-0088129P.

XX (REGC) UNIV CALIFORNIA.

XX O'Brien US;

XX WPI; 1999-229139/19.

XX N-PSDB; AAX08488.

XX Use of prosaposin receptor agonist.

XX Claim 7; Fig 2; 90pp; English.

Prosaposin is a 70kDa glycoprotein which is proteolytically processed to generate saposins A, B, C and D, all of which are similar to each other and have a similar placement of six cysteines, a glycosylation site and conserved proline residues. Prosaposin, saposin C and prosaposin derived peptides (prosaptilides), have therapeutic applications in promoting recovery after toxic, traumatic, myocardial ischemic, degenerative and inherited lesions to the peripheral and central nervous system. Prosaposin receptor agonists (PRAs) inhibit proinflammatory cytokine-induced apoptosis by activation of the Ser/Thr protein kinase Akt. Akt releases Bcl-2 and its family members, such as BAD-Bcl-2, which inhibit apoptosis. An additional mechanism which inhibit caspases, thereby component, within several minutes after binding to the receptor, PRAs block JNK activation induced by tumor necrosis factor-alpha (TNF alpha). The activation of JNK by TNF alpha is another well known mechanism for TNF alpha-induced, as well as other proinflammatory cytokine-induced apoptosis. The method can be used for inhibiting apoptosis which is caspase-mediated or induced by a proinflammatory cytokine, for example TNF alpha or interferon-gamma. It can be used for inhibiting apoptosis associated with a disorder such as e.g. rheumatoid arthritis, Crohn's disease, irritable bowel syndrome, asthma, cardiac infarction, congestive heart failure, multiple sclerosis, acute disseminated inflammatory leukoencephalitis, progressive multifocal leukoencephalitis, Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's disease, ischemic heart disease, Guillain-Barre disease, traumatic brain injury, traumatic spinal cord injury, alopecia, AIDS dementia, cerebral malaria, HIV, neuropathy, inflammatory neurodegenerative disease, and toxin-induced liver disease. This 524 N-terminal peptide of prosaposin also acts as a prosaposin receptor agonist

XX Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 2; Length 524;
 Best Local Similarity 100.0%; Pred. No. 7.2e-235;

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 LASLLGALLAGPVGLGKCTRGSAVWCQNVKTASDCGAVHCLQTVWKNKPTVKSLLPCDIC 60
DB 7 LASLLGALLAGPVGLGKCTRGSAVWCQNVKTASDCGAVHCLQTVWKNKPTVKSLLPCDIC 66
QY 61 KDVTAAAGDMLKDNATEEELIVLEKTCMDLPKPNMSASCKEIVDSYLPVILDIITIKGEMS 120
DB 67 KDVTAAAGDMLKDNATEEELIVLEKTCMDLPKPNMSASCKEIVDSYLPVILDIITIKGEMS 126
QY 121 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPEMANIPLLLYPODGPR 180
DB 127 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPEMANIPLLLYPODGPR 186
QY 181 SKPQKNDGVCDQCIQMTVDIQTAVRTNSTFVQALVEHKEECDRLGPGMADICKNYIS 240
DB 187 SKPQKNDGVCDQCIQMTVDIQTAVRTNSTFVQALVEHKEECDRLGPGMADICKNYIS 246
QY 241 QYSEIATQMMHMQPEKICALVGFCDVEKEMPQTLVPAKVASKNVVPALTEVEPIKXHE 300

```

Db 247 QYSEIATQMNMHPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIALBELVEPIKXHE 306
 QY 301 VPAKSDVYCEVCEFLVKEVTTLIDNNKTEKEILDAFDMKCSKLPSLSSECOEVDVTGS 360
 Db 307 VPASDVYCEVCEFLVKEVTTLIDNNKTEKEILDAFDMKCSKLPSLSSECOEVDVTGS 366
 QY 361 SILSLILEEVSPELVCSMLHLCSTGRLPALTVHTVTPQPDGFCFCEVCKLVGLDRLNLEKN 420
 Db 367 SILSLILEEVSPELVCSMLHLCSTGRLPALTVHTVTPQPDGFCFCEVCKLVGLDRLNLEKN 426
 QY 421 STKOEIIAALAEKGSFLPDPIYQKCDQFAVEYEPVLIEILVWMPSPFVCLIKIGACPSAH 480
 Db 427 STKOEIIAALAEKGSFLPDPIYQKCDQFAVEYEPVLIEILVWMPSPFVCLIKIGACPSAH 486
 QY 481 KPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVNN 518
 Db 487 KPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVNN 524
 RESULT 3
 AAY58716
 ID AAY58716 standard; protein, 524 AA.
 AC AAY58716;
 DT 25-APR-2000 (first entry)
 XX Human prosaposin.
 DE
 KW Prosaposin; saposin B; antiangiogenic; angiogenesis inhibitor;
 KW antitumour; antiproliferative; antimigratory; Kaposi's sarcoma; tumour;
 KW human; therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Protein
 FT /note= "mature saposin B"
 FT /note= "mature saposin B"
 FT /note= "specifically claimed antiangiogenic peptide of
 FT Claim 23"
 FT /note= "specifically claimed antiangiogenic peptide of
 FT Claim 4"
 FT Peptide
 XX
 FN WO200002902-A1.
 PD 20-JAN-2000.
 XX
 PD 12-JUL-1999; 99WO-US015772.
 XX
 PR 13-JUL-1998; 98US-0092647P.
 XX
 PA (GILL/) GILL P S.
 PI
 XX GILL PS;
 XX
 DR WPI; 2000-171128/15.
 XX
 PT Saposin B derived peptides, useful as inhibitors of angiogenesis and
 PT tumor growth.
 XX
 PS Disclosure: Page 18; 78pp; English.
 XX
 CC The present sequence is that of human prosaposin, a precursor of saposin
 CC B. The invention is based on the discovery that saposin B, previously
 CC known to be involved in the hydrolysis of sphingolipids, has potent
 CC antiangiogenic and antitumour activity, and also has antiproliferative
 CC and antimigratory activity against endothelial cells. This activity is
 CC conserved in cryptic polypeptides as small as 5 amino acids (see AAY58684
 CC -715), which can be synthetically prepared and used in vitro or in vivo
 CC for the treatment of undesired angiogenesis and tumor growth, especially
 CC Kaposi's sarcoma (claimed). The polypeptides can also be used in

CC conjunction with cytotoxic moieties to selectively kill certain cell
 CC types, e.g. for treatment of cancer, angiodioma, neovascular glaucoma,
 CC arteriovenous malformation, nonunion fracture, arthritis and other
 CC connective tissue disorders, Osler-Weber syndrome, atherosclerotic
 CC plaque, psoriasis, corneal graft neovascularization, pyogenic granuloma,
 CC retrolental fibroplasia, diabetic retinopathy, scleroderma, haemangioma,
 CC trachoma, vascular adhesions and hypertrophic scars
 XX
 SQ Sequence 524 AA;
 Query Match 100.0%; Score 2759; DB 3; Length 524;
 Best Local Similarity 100.0%; Pred. No. 7.2e-235;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LASLGAALAPVIGLKECTGSAVWCNVTTADDCGAVKCLQTWNKPRVKSIPCLIC 60
 Db 7 LASLGAALAPVIGLKECTGSAVWCNVTTADDCGAVKCLQTWNKPRVKSIPCLIC 66
 QY 61 KDVTTAAGDMLKDNATEEELVYLEKTCMDLPKPMMSASCKEIVDSYLPVLLDIKGEHS 120
 Db 67 KDVTTAAGDMLKDNATEEELVYLEKTCMDLPKPMMSASCKEIVDSYLPVLLDIKGEHS 126
 QY 121 RPEVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVAVAFPMANIPLLVPQDPR 180
 Db 127 RPEVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVAVAFPMANIPLLVPQDPR 186
 QY 181 SKPQKNGDYCCDICIQWNTDIOAVRTNSFTFVALVHVKEBCDRLGPMADICKNTIS 240
 Db 187 SKPQKNGDYCCDICIQWNTDIOAVRTNSFTFVALVHVKEBCDRLGPMADICKNTIS 246
 QY 241 QYSEIATQMNMHPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIALBELVEPIKXHE 300
 Db 247 QYSEIATQMNMHPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIALBELVEPIKXHE 306
 QY 301 VPAKSDVYCEVCEFLVKEVTTLIDNNKTEKEILDAFDMKCSKLPSLSSECOEVDVTGS 360
 Db 307 VPAKSDVYCEVCEFLVKEVTTLIDNNKTEKEILDAFDMKCSKLPSLSSECOEVDVTGS 366
 QY 361 SILSLILEEVSPELVCSMLHLCSTGRLPALTVHTVTPQPDGFCFCEVCKLVGLDRLNLEKN 420
 Db 367 SILSLILEEVSPELVCSMLHLCSTGRLPALTVHTVTPQPDGFCFCEVCKLVGLDRLNLEKN 426
 QY 421 STKOEIIAALAEKGSFLPDPIYQKCDQFAVEYEPVLIEILVWMPSPFVCLIKIGACPSAH 480
 Db 427 STKOEIIAALAEKGSFLPDPIYQKCDQFAVEYEPVLIEILVWMPSPFVCLIKIGACPSAH 486
 QY 481 KPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVNN 518
 Db 487 KPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVNN 524
 RESULT 4
 ABU79099
 ID ABU79099 standard; protein, 524 AA.
 XX
 AC ABU79099;
 XX
 DT 18-JUN-2003 (first entry)
 XX
 DE Lip-TAA binding protein, Prosaposin.
 XX
 KW Superantigen; SAg; staphylococcal enterotoxin; tumour; cancer; apoptosis;
 KW gene therapy; mammalian cell receptor; cytostatic;
 KW tumour associated lipid; energy; T cell; antigen presenting cell; APC;
 KW tumouricidal immunocyte; antitumour.
 OS Unidentified.
 XX
 FN US2002177551-A1.
 XX
 PD 28-NOV-2002.
 XX
 PF 30-MAY-2001; 2001US-00870759.

XX 31-MAY-2000; 2000US-0208128P.
 PR (TERM.) TERMAN D S.
 PA
 XX Terman DS;
 PT WPI; 2003-361759/34.
 DR
 XX A mammalian cell receptor, useful in the treatment of cancer by binding
 PT to tumor associated lipids where the binding induces anergy or apoptosis
 PT in T cells and antigen presenting cells.
 PS Disclosure; Page; 167pp; English.
 XX The invention relates to a mammalian cell receptor, useful in the
 CC treatment of cancer, which binds to tumor associated lipids and induces
 CC anergy or apoptosis in the T cells and antigen presenting cells (APCs).
 CC Also included are a mammalian cell useful in the treatment of cancer
 CC where the receptor which binds tumor associated lipids and induces
 CC cellular inactivation or death is deleted or functionally deactivated,
 CC producing (M1) a tumouricidal immunocyte population in vivo in a mammal
 CC (by allowing tumour associated lipids to contact immunocytes in which
 CC receptors for immunosuppressive fatty acids, ceramides, glycolipids,
 CC sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides,
 CC sialylated glycans, lipopeptides and proteoglycolipids are inactivated or
 CC deleted), a construct useful in the treatment of cancer comprising a
 CC superantigen (SAg) nucleotide inserted into a virus, a mammalian T cell
 CC useful in the treatment of cancer (where an adaptor protein which
 CC inhibits T cell activation by tumour associated antigens is deleted or
 CC functionally deactivated), a composition useful in the treatment of
 CC cancer (comprising a lipid raft conjugated to a superantigen), producing
 CC (M2) a tumouricidal immunocyte population ex vivo in a mammal (by
 CC allowing tumour associated lipids to contact immunocytes, in which
 CC receptors for the lipids are inactivated or deleted to produce a
 CC tumouricidal immunocyte population, and administering the tumouricidal
 CC population ex vivo in a mammal (by allowing a tumour associated lipid to
 CC contact APCs, in which receptors for the tumour associated lipids are
 CC inactivated or deleted to produce a tumouricidally activated population,
 CC and administering APCs to the host), producing a tumouricidal T cell
 CC population ex vivo in a mammal (by allowing a tumour associated lipid to
 CC contact T cells, in which adaptor proteins, which inhibit T cell
 CC activation by tumour associated antigens, are deleted or functionally
 CC deactivated to produce a tumouricidal population of T cells, and
 CC administering the tumouricidally activated T cells to the host, or
 CC allowing a superantigen-lipid raft to contact T cells ex vivo, and
 CC administering the tumouricidally activated T cells to the host), treating
 CC (M5) cancer in a mammal (by administering a lipid binding molecule which
 CC binds immunosuppressive tumour associated lipids in vivo), producing (M6)
 CC a tumouricidal T cell population in vivo in a mammal (by allowing a
 CC tumour associated antigen to contact immunocytes in which adaptor
 CC proteins which inhibit T cell activation by tumour associated antigens
 CC are deleted or functionally deactivated) and producing (M7) a
 CC tumouricidal T cell population ex vivo in a mammal comprising allowing a
 CC superantigen-lipid raft conjugate to contact immunocytes in vivo. The
 CC receptors, methods and compositions are useful for treating cancers and
 CC tumours. Bacterial superantigens are co-administered or administered as
 CC fusion constructs with anti-tumour proteins or motifs. The present
 CC sequence represents a tumour antigen or a motif identifying a tumour
 CC antigen, which can be functionally deactivated in the method of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format from the
 CC US patent office website at
 CC "seqdata.uspto.gov/sequence.html?DocID=20020177551"
 CC
 XX Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 7, 2e-235;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 LASLIGALLAPVILGKECTRGSAVWCQNVKTASDCGAVKHLQTVWNRKPTVSLPCDIC 60

DB 7 LASLIGALLAPVILGKECTRGSAVWCQNVKTASDCGAVKHLQTVWNRKPTVSLPCDIC 66
 QY 61 KDVTYTAAGDMLKDNATEBEILLVLEKTCDWLPKRNMSASCKEIVDSYLPVILDIKGEWS 120
 DB 67 KDVTYTAAGDMLKDNATEBEILLVLEKTCDWLPKRNMSASCKEIVDSYLPVILDIKGEWS 126
 QY 121 RPEVCSALNLCESLQRLAEILNHQKLESNKI PELUMTEVVA PPMANTILLYPQDGR 180
 DB 127 RPEVCSALNLCESLQRLAEILNHQKLESNKI PELUMTEVVA PPMANTILLYPQDGR 186
 QY 181 SKPQKNDGVCDQCIQMTVDIQTAVRTNSTFQALVEHVKESCDRLGPMADICKNYIS 240
 DB 187 SKPQKNDGVCDQCIQMTVDIQTAVRTNSTFQALVEHVKESCDRLGPMADICKNYIS 246
 QY 241 QYSEIATQMMHMQPKICALVGPCEVEMPMQTLVPAKVASKNYIPALTEVPIKKE 300
 DB 247 QYSEIATQMMHMQPKICALVGPCEVEMPMQTLVPAKVASKNYIPALTEVPIKKE 306
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSKLPKSLSECCQVVDYGS 360
 DB 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSKLPKSLSECCQVVDYGS 366
 QY 361 SIISILIEEVSPELVCSMLHLCSTGLPALTYHVTQPKQGGCEVCKLVGYLDRLERN 420
 DB 367 SIISILIEEVSPELVCSMLHLCSTGLPALTYHVTQPKQGGCEVCKLVGYLDRLERN 426
 QY 421 STQKEILALKEKCSFLPPYQKQCFVAEYEPVLEILVEVMDPSFVCLTIGACPSAH 480
 DB 427 STQKEILALKEKCSFLPPYQKQCFVAEYEPVLEILVEVMDPSFVCLTIGACPSAH 486
 QY 481 KPILGTEKICWGSYWCQNTETPAQCNVAEHCKRHVYN 518
 DB 487 KPILGTEKICWGSYWCQNTETPAQCNVAEHCKRHVYN 524

RESULT 5
 ABU05200
 ID ABU05200 standard; protein; 524 AA.
 XX
 AC ABU05200;
 XX
 XX 29-JAN-2003 (first entry)
 DE Human expressed protein tag (EPT) #1866.
 XX
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200278524-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 28-MAR-2002; 2002WO-US009671.
 XX
 PR 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 PA (ZYCO-) ZYCO INC.
 XX
 PI Chicx RM, Tomlinson AJ, Urban RG;
 XX WPI; 2003-040607/03.

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.

Example 2; SEQ ID NO 1866; 134bp; English.

The invention describes a purified polypeptide, which comprises a
fragment of a kinase, phosphatase, protease, protease inhibitor,
transporter, cytoskeletal protein, receptor or transcription factor. The
polypeptide is useful as an immunogenic composition for eliciting in a
mammal an immunogenic response directed against any of the purified
polypeptide. The purified polypeptide, or the antibody that binds to this
polypeptide, is useful for treating cancer. The polypeptide is also
useful for identifying compounds that binds to a naturally processed
class I or class II MHC-binding polypeptide. The polypeptides and
polynucleotides are particularly useful for treating or preventing
myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
lymphoma or leukemia. These are also useful for screening agents for
treating the above mentioned diseases. This sequence represents an
expressed protein tag (EPT) isolated from human tissue for translational
profiling. Note: This sequence does not appear in the printed
specification but was obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 7.2e-235;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
1 LASLLGALAGPVGLKECTRGSAVWCONVKTASDCGAVKICLOTVMNKPVTYKSLPCDIC 60
7 LASLLGALAGPVGLKECTRGSAVWCONVKTASDCGAVKICLOTVMNKPVTYKSLPCDIC 66
61 KDVTAAADMLKDNTAEELIVLEKTCMDMLPKPMMSASCKEIVSYLVILDIKIGMS 120
67 KDVTAAADMLKDNTAEELIVLEKTCMDMLPKPMMSASCKEIVSYLVILDIKIGMS 126
121 RPEVCSALNLCESIQKHLAEINHQKLESNKIPELDTEVAVAPMANIPLLVPQDGR 180
127 RPEVCSALNLCESIQKHLAEINHQKLESNKIPELDTEVAVAPMANIPLLVPQDGR 186
181 SKPQKNDGVCCDICTQVAVTDTQVAVRNSFTVQALVEHVKEECRLGPGMADICKNTS 240
187 SKPQKNDGVCCDICTQVAVTDTQVAVRNSFTVQALVEHVKEECRLGPGMADICKNTS 246
241 QYSEIAIQMMHMQKEICAVGFCEVEMPMQTLVPAKVASKNVIRPALVEPIKXHE 300
247 QYSEIAIQMMHMQKEICAVGFCEVEMPMQTLVPAKVASKNVIRPALVEPIKXHE 306
301 VPASDVYCEVCEPLVKEVTXKIDNNKTEKEILADPKKCSKLPKSLSEBCEQVVDYGS 360
307 VPASDVYCEVCEPLVKEVTXKIDNNKTEKEILADPKKCSKLPKSLSEBCEQVVDYGS 366
361 SILSLLEEVSPRLVCSMLHICSGTRLPALTVAHVQPKDGGCEVCKLGVLDNLEKN 420
367 SILSLLEEVSPRLVCSMLHICSGTRLPALTVAHVQPKDGGCEVCKLGVLDNLEKN 426
421 STKOEIIAALAEKGCFLPDYVQKQCDQFAVEYEPVLIETLVMPSPFVCLTIGACPSAH 480
427 STKOEIIAALAEKGCFLPDYVQKQCDQFAVEYEPVLIETLVMPSPFVCLTIGACPSAH 486
481 KPILGTEKCIWGPSTWCONTEFAAOCAVHECKRHVN 518
487 KPILGTEKCIWGPSTWCONTEFAAOCAVHECKRHVN 524
```

RESULT 6
ABU05207
ID ABU05207 standard; protein; 524 AA.
XX
AC ABU05207;

XX 29-JAN-2003 (first entry)
DT
XX
DE Human expressed protein tag (EPT) #1873.

KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.

OS Homo sapiens.

MO200278524-A2.

10-OCT-2002.

28-MAR-2002; 2002WO-US009671.

28-MAR-2001; 2001US-0279495P.

21-MAY-2001; 2001US-0292544P.

08-AUG-2001; 2001US-0310801P.

01-OCT-2001; 2001US-0326370P.

04-DEC-2001; 2001US-0336780P.

20-FEB-2002; 2002US-0358985P.

(ZYCO-) ZYCO INC.

Chicz RM, Tomlinson AJ, Urban RG;

WPI; 2003-040607/03.

Example 2; SEQ ID NO 1873; 134bp; English.

The invention describes a purified polypeptide, which comprises a
fragment of a kinase, phosphatase, protease, protease inhibitor,
transporter, cytoskeletal protein, receptor or transcription factor. The
polypeptide is useful as an immunogenic composition for eliciting in a
mammal an immunogenic response directed against any of the purified
polypeptide. The purified polypeptide, or the antibody that binds to this
polypeptide, is useful for treating cancer. The polypeptide is also
useful for identifying compounds that binds to a naturally processed
class I or class II MHC-binding polypeptide. The polypeptides and
polynucleotides are particularly useful for treating or preventing
myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
lymphoma or leukemia. These are also useful for screening agents for
treating the above mentioned diseases. This sequence represents an
expressed protein tag (EPT) isolated from human tissue for translational
profiling. Note: This sequence does not appear in the printed
specification but was obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 7.2e-235;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
1 LASLLGALAGPVGLKECTRGSAVWCONVKTASDCGAVKICLOTVMNKPVTYKSLPCDIC 60
7 LASLLGALAGPVGLKECTRGSAVWCONVKTASDCGAVKICLOTVMNKPVTYKSLPCDIC 66
61 KDVTAAADMLKDNTAEELIVLEKTCMDMLPKPMMSASCKEIVSYLVILDIKIGMS 120
67 KDVTAAADMLKDNTAEELIVLEKTCMDMLPKPMMSASCKEIVSYLVILDIKIGMS 126
121 RPEVCSALNLCESIQKHLAEINHQKLESNKIPELDTEVAVAPMANIPLLVPQDGR 180
```

Db 127 RPEVCSALNLCESLQKHLAEINHQKLESNKIPBELDMTEVVAEPMANIPILLYPQDGP 186
 QY 181 SKPQKNDGVDCDCIQWVTDIQTAVRTNSTFVQALVEHYVEECRLPGMADICKNYIS 240
 Db 187 SKPQKNDGVDCDCIQWVTDIQTAVRTNSTFVQALVEHYVEECRLPGMADICKNYIS 246
 QY 241 QYSEIATQMMHMOPEKICALVGFCDVEKEMQTLVPAVASKNVLPALVEPIKKEH 300
 Db 247 QYSEIATQMMHMOPEKICALVGFCDVEKEMQTLVPAVASKNVLPALVEPIKKEH 306
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDTYGS 360
 Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDTYGS 366
 QY 361 SILSILIEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKLVGYIDRNLEKN 420
 Db 367 SILSILIEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKLVGYIDRNLEKN 426
 QY 421 STKQEIILAEKGCSEFLPDYQKQCDQFVAEYEPVLIELIVEVMDPSFVCLIKGACPSAH 480
 Db 427 STKQEIILAEKGCSEFLPDYQKQCDQFVAEYEPVLIELIVEVMDPSFVCLIKGACPSAH 486
 QY 481 KPLGTGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVWN 518
 Db 487 KPLGTGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVWN 524

RESULT 7

ABU05203
 ID ABU05203 standard; protein; 524 AA.

ABU05203;

29-JAN-2003 (first entry)

Human expressed protein tag (EPT) #1869.

Translation profiling; expressed protein tag; EPT; kinase; phosphatase;
 protease; protease inhibitor; transporter; cytoskeletal protein;
 receptor; transcription factor; cancer; MHC;
 major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.

Homo sapiens.

MO200278524-A2.

10-OCT-2002.

28-MAR-2002; 2002MO-US009671.

28-MAR-2001; 2001US-0279495P.

21-MAY-2001; 2001US-0292544P.

08-AUG-2001; 2001US-0310801P.

01-OCT-2001; 2001US-0326370P.

04-DEC-2001; 2001US-0336780P.

20-FEB-2002; 2002US-0358985P.

(ZYCO-) ZYCOS INC.

Chicz RM, Tomlinson AJ, Urban RG;

WPI; 2003-040607/03.

New polypeptides (e.g. kinases, phosphatases, proteases, transporters,

cytoskeletal proteins, receptors or transcription factors), useful for

treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or

leukemia.

Example 2; SEQ ID NO 1869; 134pp; English.

CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that bind to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polymulectins are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIP0 at
 CC ftp.wipo.int/pub/published_pct_sequences

Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 7.2e-235; Mismatches 0; Indels 0; Gaps 0;

Db 1 LASTLGAALGAPVLGLKECTRGSAVWCQNVKTASDCAVGHCLQTVWKNPKYSIPCDIC 60
 QY 7 LASTLGAALGAPVLGLKECTRGSAVWCQNVKTASDCAVGHCLQTVWKNPKYSIPCDIC 66
 Db 61 KDVTAAAGDMLKONATEEELVLEKTCMDLPRPNNSASCKEIVDSYLPVILLDIKGEWS 120
 QY 67 KDVTAAAGDMLKONATEEELVLEKTCMDLPRPNNSASCKEIVDSYLPVILLDIKGEWS 126
 Db 121 RPEVCSALNLCESLQKHLAEINHQKLESNKIPBELDMTEVVAEPMANIPILLYPQDGP 180
 QY 127 RPEVCSALNLCESLQKHLAEINHQKLESNKIPBELDMTEVVAEPMANIPILLYPQDGP 186
 Db 181 SKPQKNDGVDCDCIQWVTDIQTAVRTNSTFVQALVEHYVEECRLPGMADICKNYIS 240
 QY 187 SKPQKNDGVDCDCIQWVTDIQTAVRTNSTFVQALVEHYVEECRLPGMADICKNYIS 246
 Db 241 QYSEIATQMMHMOPEKICALVGFCDVEKEMQTLVPAVASKNVLPALVEPIKKEH 300
 QY 247 QYSEIATQMMHMOPEKICALVGFCDVEKEMQTLVPAVASKNVLPALVEPIKKEH 306
 Db 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDTYGS 360
 QY 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDTYGS 366
 Db 361 SILSILIEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKLVGYIDRNLEKN 420
 QY 367 SILSILIEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKLVGYIDRNLEKN 426
 Db 421 STKQEIILAEKGCSEFLPDYQKQCDQFVAEYEPVLIELIVEVMDPSFVCLIKGACPSAH 480
 QY 427 STKQEIILAEKGCSEFLPDYQKQCDQFVAEYEPVLIELIVEVMDPSFVCLIKGACPSAH 486
 Db 481 KPLGTGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVWN 518
 QY 487 KPLGTGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVWN 524

RESULT 8

ABU07340

ID ABU07340 standard; protein; 524 AA.

ABU07340;

29-JAN-2003 (first entry)

Human expressed protein tag (EPT) #2041.

Translation profiling; expressed protein tag; EPT; kinase; phosphatase;
 protease; protease inhibitor; transporter; cytoskeletal protein;
 receptor; transcription factor; cancer; MHC;
 major histocompatibility complex; myeloma; colon cancer; gastric cancer;


```

KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Chicx RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
PS Example 2; SEQ ID NO 2041; 134pp; English.
XX
CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 524 AA:
Query Match 100.0%; Score 2759; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 7.2e-235;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 301 VPAKSDVYCEVCEFLVKEVNTGLIDNNKTEKEIILDAFDKWCSEKLSSEGOEVDVYGS 360
DB 307 VPAKSDVYCEVCEFLVKEVNTGLIDNNKTEKEIILDAFDKWCSEKLSSEGOEVDVYGS 366
QY 361 SILSILLEEVSPELVCSMLHCSGSTRLPALVHTVTPKDGSCFEVCKLVGLDRLNLEKN 420
DB 367 SILSILLEEVSPELVCSMLHCSGSTRLPALVHTVTPKDGSCFEVCKLVGLDRLNLEKN 426
QY 421 STKOEIILAEKSGSFLLPDPYOKOQDFVAEYEVLLIILVWVNDPSVCLKIGACPSAH 480
DB 427 STKOEIILAEKSGSFLLPDPYOKOQDFVAEYEVLLIILVWVNDPSVCLKIGACPSAH 486
QY 481 KPLIGTEKCIWGPSTWCONTEETAACNAVEHCKRHWN 518
DB 487 KPLIGTEKCIWGPSTWCONTEETAACNAVEHCKRHWN 524

RESULT 9
ABU05216
ID ABU05216 standard; protein; 524 AA.
XX
AC ABU05216;
XX
DT 29-JAN-2003 (first entry)
XX
DB Human expressed protein tag (EPT) #1882.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Chicx RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
PS Example 2; SEQ ID NO 1882; 134pp; English.
XX
CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,

```

CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 7.2e-235;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLIGALAGVILGKECTRGSAVWQNVKTASDGAHVHCLQTVMNKPVTYKSLPCDIC 60
 DB 7 LASLIGALAGVILGKECTRGSAVWQNVKTASDGAHVHCLQTVMNKPVTYKSLPCDIC 66
 QY 61 KDVTAAAGDMLKDNATBEELIYVLEKTCMDLPRKMSASCKEIVDSYLPVLDIYKEMS 120
 DB 67 KDVTAAAGDMLKDNATBEELIYVLEKTCMDLPRKMSASCKEIVDSYLPVLDIYKEMS 126
 QY 121 RPEVCSALNLCESLQKHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLIYPODGR 180
 DB 127 RPEVCSALNLCESLQKHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLIYPODGR 186
 QY 181 SKQPKNDGVCCODCIQWVTDIQTAVRNSTFVQALVEHKEBCDRIQGMADICKNYIS 240
 DB 187 SKQPKNDGVCCODCIQWVTDIQTAVRNSTFVQALVEHKEBCDRIQGMADICKNYIS 246
 QY 241 QYSEIADQMMHMQPEKICALVGFCDVEKEMQTLVPAKVASKNVIALBELVEPIKHE 300
 DB 247 QYSEIADQMMHMQPEKICALVGFCDVEKEMQTLVPAKVASKNVIALBELVEPIKHE 306
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSECOEVDYTG 360
 DB 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSECOEVDYTG 366
 QY 361 SILSILIEEVSPELVCSMLHLCSTGRPLPALVHTVTPKDGFCVEVCKLVGYLDRLNEKN 420
 DB 367 SILSILIEEVSPELVCSMLHLCSTGRPLPALVHTVTPKDGFCVEVCKLVGYLDRLNEKN 426
 QY 421 STKOEIILAEKGCSEFLPDYPOKODQVFAEYEPVLEIIVWMDPSVCLKIGACPSAH 480
 DB 427 STKOEIILAEKGCSEFLPDYPOKODQVFAEYEPVLEIIVWMDPSVCLKIGACPSAH 486
 QY 481 KPLLTGKTCIWPSTYWCNTETAAQCNAVEHCKRHVN 518
 DB 487 KPLLTGKTCIWPSTYWCNTETAAQCNAVEHCKRHVN 524

RESULT 10
 ABU05202
 ID ABU05202 standard; protein; 524 AA.

AC ABU05202;

XX 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #1868.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

OS Homo sapiens.

XX WO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 06-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCO- INC.

XX Chicz RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.

XX Example 2; SEQ ID NO 1868; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide, is useful for treating cancer. The polypeptide binds to this
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC lymphoma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 7.2e-235;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLIGALAGVILGKECTRGSAVWQNVKTASDGAHVHCLQTVMNKPVTYKSLPCDIC 60
 DB 7 LASLIGALAGVILGKECTRGSAVWQNVKTASDGAHVHCLQTVMNKPVTYKSLPCDIC 66
 QY 61 KDVTAAAGDMLKDNATBEELIYVLEKTCMDLPRKMSASCKEIVDSYLPVLDIYKEMS 120
 DB 67 KDVTAAAGDMLKDNATBEELIYVLEKTCMDLPRKMSASCKEIVDSYLPVLDIYKEMS 126
 QY 121 RPEVCSALNLCESLQKHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLIYPODGR 180
 DB 127 RPEVCSALNLCESLQKHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLIYPODGR 186
 QY 181 SKQPKNDGVCCODCIQWVTDIQTAVRNSTFVQALVEHKEBCDRIQGMADICKNYIS 240
 DB 187 SKQPKNDGVCCODCIQWVTDIQTAVRNSTFVQALVEHKEBCDRIQGMADICKNYIS 246
 QY 241 QYSEIADQMMHMQPEKICALVGFCDVEKEMQTLVPAKVASKNVIALBELVEPIKHE 300
 DB 247 QYSEIADQMMHMQPEKICALVGFCDVEKEMQTLVPAKVASKNVIALBELVEPIKHE 306
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSECOEVDYTG 360
 DB 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSECOEVDYTG 366
 QY 361 SILSILIEEVSPELVCSMLHLCSTGRPLPALVHTVTPKDGFCVEVCKLVGYLDRLNEKN 420
 DB 367 SILSILIEEVSPELVCSMLHLCSTGRPLPALVHTVTPKDGFCVEVCKLVGYLDRLNEKN 426

QY 421 STKQETLLAALKGCSFLPDPYQKQCDQFAVEYEPVLIILVEWMDPSFVCLKIGACPSAH 480
 DB 427 STKQETLLAALKGCSFLPDPYQKQCDQFAVEYEPVLIILVEWMDPSFVCLKIGACPSAH 486
 QY 481 KPLIGTEKCIWGPSYWCNTETTAACNAVEHCKRHVMN 518
 DB 487 KPLIGTEKCIWGPSYWCNTETTAACNAVEHCKRHVMN 524

RESULT 11
 ABU05214
 ID ABU05214 standard; protein; 524 AA.
 AC ABU05214;
 XX
 DT 29-JAN-2003 (first entry)
 XX
 DE Human expressed protein tag (EPT) #1880.
 XX
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 XX
 OS Homo sapiens.
 XX
 EN WO200278524-A2.
 XX
 PD 10-OCT-2002.
 XX
 PE 28-MAR-2002; 2002MO-US009671.
 XX
 PR 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 PA (ZYCO-) ZYCOS INC.
 XX
 PI Chicz RM, Tomlinson AJ, Urban RG;
 XX
 DR WPI; 2003-040607/03.
 XX
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX
 PS Example 2; SEQ ID NO 1880; 134pp; English.
 XX
 CC The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukaemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIP0 at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 6; Length 524;
 Best Local Similarity 100.0%; Pred. No. 7, 2e-235;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLIGALAPVIGLKEKTRGSAAVWQNVKTASDGAVKQLQTVWKKPTVKSIPCDIC 60
 DB 7 LASLIGALAPVIGLKEKTRGSAAVWQNVKTASDGAVKQLQTVWKKPTVKSIPCDIC 66
 QY 61 KDVTAAADMDKDNATEEELIIVLEKTCMDLPKPMMSASCKEIVDSYLPVLLDIKIMS 120
 DB 67 KDVTAAADMDKDNATEEELIIVLEKTCMDLPKPMMSASCKEIVDSYLPVLLDIKIMS 126
 QY 121 RPEVCSALNICESLQKILAEINHQKLESNKIPELDTEVVAPEPMANIPLLTYQDGR 180
 DB 127 RPEVCSALNICESLQKILAEINHQKLESNKIPELDTEVVAPEPMANIPLLTYQDGR 186
 QY 181 SKPQKNDGVCCDICIQWTDIQTAVRTNSTFVQALVEHVKECDRLGPMADICKNYIS 240
 DB 187 SKPQKNDGVCCDICIQWTDIQTAVRTNSTFVQALVEHVKECDRLGPMADICKNYIS 246
 QY 241 QYSEIATQMMHMQPKETCALVGFCDVEKEMPMQTLVPAKVASKNVITALEIVEPIKGE 300
 DB 247 QYSEIATQMMHMQPKETCALVGFCDVEKEMPMQTLVPAKVASKNVITALEIVEPIKGE 306
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDGKCSKLPKSLSECCQEVDTYGS 360
 DB 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDGKCSKLPKSLSECCQEVDTYGS 366
 QY 361 SILSLILEEVSPELVCSMLHLCSGTRLPALTIVHTQPRDGGCEVCKLVGLDNLNLEKN 420
 DB 367 SILSLILEEVSPELVCSMLHLCSGTRLPALTIVHTQPRDGGCEVCKLVGLDNLNLEKN 426
 QY 421 STKQETLLAALKGCSFLPDPYQKQCDQFAVEYEPVLIILVEWMDPSFVCLKIGACPSAH 480
 DB 427 STKQETLLAALKGCSFLPDPYQKQCDQFAVEYEPVLIILVEWMDPSFVCLKIGACPSAH 486
 QY 481 KPLIGTEKCIWGPSYWCNTETTAACNAVEHCKRHVMN 518
 DB 487 KPLIGTEKCIWGPSYWCNTETTAACNAVEHCKRHVMN 524

RESULT 12
 ABU05215
 ID ABU05215 standard; protein; 524 AA.
 AC ABU05215;
 XX
 DT 29-JAN-2003 (first entry)
 XX
 DE Human expressed protein tag (EPT) #1881.
 XX
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
 KW protease; protease inhibitor; transporter; cytoskeletal protein;
 KW receptor; transcription factor; cancer; MHC;
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
 XX
 OS Homo sapiens.
 XX
 EN WO200278524-A2.
 XX
 PD 10-OCT-2002.
 XX
 PE 28-MAR-2002; 2002MO-US009671.
 XX
 PR 28-MAR-2001; 2001US-0279495P.
 PR 21-MAY-2001; 2001US-0292544P.
 PR 08-AUG-2001; 2001US-0310801P.
 PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 PA (ZYCO-) ZYCOS INC.

XX Chicz RM, Tomlinson AJ, Urban RG;
XX
XX WPI, 2003-040607/03.
DR

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.

PS Example 2; SEQ ID NO 1881; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptide and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 7.2e-235;

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 LASLIGALAGPVILGKECTRGSAVWCQNVKTAASDCGAVGHCLOTWNKPTVKSIPCDIC 60
DB 7 LASLIGALAGPVILGKECTRGSAVWCQNVKTAASDCGAVGHCLOTWNKPTVKSIPCDIC 66
QY 61 KDVVTAAGDMLKDNATTEELIVLYLEKTCDWLPKPNMSASCKEIVDSYLPVLIIDIIKEMS 120
DB 67 KDVVTAAGDMLKDNATTEELIVLYLEKTCDWLPKPNMSASCKEIVDSYLPVLIIDIIKEMS 126
QY 121 RGEVCSALNLCESLQKHLAEINHOKESENKIPBLDTEVVAAPMANIPILLIYPQDGP 180
DB 127 RGEVCSALNLCESLQKHLAEINHOKESENKIPBLDTEVVAAPMANIPILLIYPQDGP 186
QY 181 SKPQKNDGVQCQDCIQWVTDIQTAVRTNSTFEVQALVHVKEECDRIGPGWADICKNYIS 240
DB 187 SKPQKNDGVQCQDCIQWVTDIQTAVRTNSTFEVQALVHVKEECDRIGPGWADICKNYIS 246
QY 241 QYSEAIQMMHMPKKEICAVGFCDEKVEKPMQTLVPKAKASNVIPALEVEPIKHE 300
DB 247 QYSEAIQMMHMPKKEICAVGFCDEKVEKPMQTLVPKAKASNVIPALEVEPIKHE 306
QY 301 VPAASDVYCEVCEFLVKEVTKLINNKTEKEIILAFPMKSKIPKSISEECQEVPTGYS 360
DB 307 VPAASDVYCEVCEFLVKEVTKLINNKTEKEIILAFPMKSKIPKSISEECQEVPTGYS 366
QY 361 SIISILIEEVSPELVCSMLHICSGTRLPALTVAHTQPKDGFCEVCKKLGVYLDRLNLEK 420
DB 367 SIISILIEEVSPELVCSMLHICSGTRLPALTVAHTQPKDGFCEVCKKLGVYLDRLNLEK 426
QY 421 STKOELIAALEKGSFIPDPYKOCDOFVAEYEVULLEILVEWDSFVCLKIGACPSAH 480
DB 427 STKOELIAALEKGSFIPDPYKOCDOFVAEYEVULLEILVEWDSFVCLKIGACPSAH 486
QY 481 KPLILGTEKCIWGSYWCQNTETAACQNAVEHCKRHWN 518
DB 487 KPLILGTEKCIWGSYWCQNTETAACQNAVEHCKRHWN 524

```

RESULT 13
ABU05199
ID ABU05199 standard; protein; 524 AA.
XX

XX AC ABU05199;
XX
XX DT 29-JAN-2003 (first entry)
XX

XX Human expressed protein tag (EPT) #1865.

XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.

XX Homo sapiens.

XX PN WO200278524-A2.
XX

XX PD 10-OCT-2002.
XX

XX PF 28-MAR-2002; 2002MO-US009671.
XX

XX PR 28-MAR-2001; 2001US-0279495P.
XX PR 21-MAY-2001; 2001US-0292544P.
XX PR 08-AUG-2001; 2001US-0310801P.
XX PR 01-OCT-2001; 2001US-0326370P.
XX PR 04-DEC-2001; 2001US-0336780P.
XX PR 20-FEB-2002; 2002US-0358985P.

XX PA (ZYCO-) ZYCOS INC.

XX PI Chicz RM, Tomlinson AJ, Urban RG;
XX

XX DR WPI, 2003-040607/03.

XX PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.

XX PS Example 2; SEQ ID NO 1865; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptide and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 7.2e-235;

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 LASLIGALAGPVILGKECTRGSAVWCQNVKTAASDCGAVGHCLOTWNKPTVKSIPCDIC 60
DB 7 LASLIGALAGPVILGKECTRGSAVWCQNVKTAASDCGAVGHCLOTWNKPTVKSIPCDIC 66
QY 61 KDVVTAAGDMLKDNATTEELIVLYLEKTCDWLPKPNMSASCKEIVDSYLPVLIIDIIKEMS 120

```

```
DB 67 KDVTAAAGDMLKDNATBEELIIVLEKTCQWMLPKPMNSASCKEIVDSYLPVLLIDIIKGEMS 126
QY 121 RRGVCSALNLCESIQKHLAEINHOKELESNKIPBLDMTEVVAAPPMANIPLLIYPODGR 180
DB 127 RRGVCSALNLCESIQKHLAEINHOKELESNKIPBLDMTEVVAAPPMANIPLLIYPODGR 186
QY 181 SKPOKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVHVHVEECRLGPMADICKNYIS 240
DB 187 SKPOKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVHVHVEECRLGPMADICKNYIS 246
QY 241 QYSEIAIQMMHMQKEICALVGFCDVEKEMQTLVPAKVASKNVIPAELVEPIKKE 300
DB 247 QYSEIAIQMMHMQKEICALVGFCDVEKEMQTLVPAKVASKNVIPAELVEPIKKE 306
QY 301 VPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKMSCKLPKSLSECCOEVVDYGS 360
DB 307 VPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKMSCKLPKSLSECCOEVVDYGS 366
QY 361 SILSTILLEEVSPBELVCSMLHLCSGTRLPALTVHVTQPKDGGFCVCKLVGYLDNLEKN 420
DB 367 SILSTILLEEVSPBELVCSMLHLCSGTRLPALTVHVTQPKDGGFCVCKLVGYLDNLEKN 426
QY 421 STKOEIIAALBKSGSFLPDPYQKCDQFVAEYEPVLIIEVWMPSPFVCLIKIGACPSAH 480
DB 427 STKOEIIAALBKSGSFLPDPYQKCDQFVAEYEPVLIIEVWMPSPFVCLIKIGACPSAH 486
QY 481 KPILGTEKICWGPSYWCNTETAAQCNVHCKRHVNN 518
DB 487 KPILGTEKICWGPSYWCNTETAAQCNVHCKRHVNN 524

RESULT 14
ABU05212 ID ABU05212 standard; protein; 524 AA.
AC ABU05212;
XX
XX 29-JAN-2003 (first entry)
DT
XX Human expressed protein tag (EPT) #1878.
DE
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.
XX
XX Homo sapiens.
XX
XX WO200278524-A2.
XX
XX 10-OCT-2002.
XX
XX 28-MAR-2002; 2002WO-US009671.
XX
XX 28-MAR-2001; 2001US-0279495P.
XX 21-MAY-2001; 2001US-0292544P.
XX 08-AUG-2001; 2001US-0310801P.
XX 01-OCT-2001; 2001US-0326370P.
XX 04-DEC-2001; 2001US-0336780P.
XX 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCO INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI, 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX
```

```
XX
PS Example 2; SEQ ID NO 1878; 134pp; English.
XX
CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIP0 at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 524 AA;
XX
XX Query Match 100.0%; Score 2759; DB 6; Length 524;
XX Best Local Similarity 100.0%; Pred. No. 7,2e-235;
XX Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 LASLIGAAIAGVIGLKECTRGSAVWQNVKTASDCGAVKHCLQTVNNKPTVKSIPCDIC 60
DB 7 LASLIGAAIAGVIGLKECTRGSAVWQNVKTASDCGAVKHCLQTVNNKPTVKSIPCDIC 66
QY 61 KDVTAAAGDMLKDNATBEELIIVLEKTCQWMLPKPMNSASCKEIVDSYLPVLLIDIIKGEMS 120
DB 67 KDVTAAAGDMLKDNATBEELIIVLEKTCQWMLPKPMNSASCKEIVDSYLPVLLIDIIKGEMS 126
QY 121 RRGVCSALNLCESIQKHLAEINHOKELESNKIPBLDMTEVVAAPPMANIPLLIYPODGR 180
DB 127 RRGVCSALNLCESIQKHLAEINHOKELESNKIPBLDMTEVVAAPPMANIPLLIYPODGR 186
QY 181 SKPOKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVHVHVEECRLGPMADICKNYIS 240
DB 187 SKPOKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVHVHVEECRLGPMADICKNYIS 246
QY 241 QYSEIAIQMMHMQKEICALVGFCDVEKEMQTLVPAKVASKNVIPAELVEPIKKE 300
DB 247 QYSEIAIQMMHMQKEICALVGFCDVEKEMQTLVPAKVASKNVIPAELVEPIKKE 306
QY 301 VPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKMSCKLPKSLSECCOEVVDYGS 360
DB 307 VPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKMSCKLPKSLSECCOEVVDYGS 366
QY 361 SILSTILLEEVSPBELVCSMLHLCSGTRLPALTVHVTQPKDGGFCVCKLVGYLDNLEKN 420
DB 367 SILSTILLEEVSPBELVCSMLHLCSGTRLPALTVHVTQPKDGGFCVCKLVGYLDNLEKN 426
QY 421 STKOEIIAALBKSGSFLPDPYQKCDQFVAEYEPVLIIEVWMPSPFVCLIKIGACPSAH 480
DB 427 STKOEIIAALBKSGSFLPDPYQKCDQFVAEYEPVLIIEVWMPSPFVCLIKIGACPSAH 486
QY 481 KPILGTEKICWGPSYWCNTETAAQCNVHCKRHVNN 518
DB 487 KPILGTEKICWGPSYWCNTETAAQCNVHCKRHVNN 524

RESULT 15
ABU05213 ID ABU05213 standard; protein; 524 AA.
AC ABU05213;
XX
XX 29-JAN-2003 (first entry)
DT
XX Human expressed protein tag (EPT) #1879.
XX
XX
```

XX Translational profiling: expressed protein tag; EPT, kinase; phosphatase;
KM protease; protease inhibitor; transporter; cytoskeletal protein;
KM receptor; transcription factor; cancer; MHC;
KM major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KM adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
OS Homo sapiens.
XX
XX W0200278524-A2.
XX
XX 10-OCT-2002.
XX
XX 28-MAR-2002; 2002MO-US009671.
XX
XX 28-MAR-2001; 2001US-0279495P.
XX 21-MAY-2001; 2001US-0292544P.
XX 08-AUG-2001; 2001US-0310801P.
XX 01-OCT-2001; 2001US-0326370P.
XX 04-DEC-2001; 2001US-0336780P.
XX 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCO INC.
XX
XX Chicz RM, Tomlinson AJ, Urban RG;
PI
XX
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1879; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 7,2e-235;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLLGAAAGPVYGLKRETRGSAVWCQNVYASDCAVHKLQTVWNRKPTVKSLPDIDC 60
DB 7 LASLLGAAAGPVYGLKRETRGSAVWCQNVYASDCAVHKLQTVWNRKPTVKSLPDIDC 66
QY 61 KDVTYAGDMLKDNATEEELIVLEKTCMDLPKRNMSASCKEIVDSYLPVILDIITIGEMS 120
DB 67 KDVTYAGDMLKDNATEEELIVLEKTCMDLPKRNMSASCKEIVDSYLPVILDIITIGEMS 126
QY 121 RFGVCSALNLCESLQGHIALNLHQQLBSNKLPELDMTEVAPFMANIPILLYPODGR 180
DB 127 RFGVCSALNLCESLQGHIALNLHQQLBSNKLPELDMTEVAPFMANIPILLYPODGR 186
QY 181 SKPOKNDGNDVCCQDCIQWVTDIQTAVRTNSTFVALVHVEKCDRLPGMADICKNYIS 240
|||||

DB 187 SKPOKNDGNDVCCQDCIQWVTDIQTAVRTNSTFVALVHVEKCDRLPGMADICKNYIS 246
QY 241 QYSEIAIOMMMHMQPKETCALVGFCDYKEMPMQTLVPAYASKNVTPALEVPIKHE 300
DB 247 QYSEIAIOMMMHMQPKETCALVGFCDYKEMPMQTLVPAYASKNVTPALEVPIKHE 306
QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIILDAFDMKCSKLPKSLSECCQEVVDYGS 360
DB 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIILDAFDMKCSKLPKSLSECCQEVVDYGS 366
QY 361 SILSILIEVSPPELVCSMLHCSGTRLPALTTHVTOPKDGCFCEVCKKLGVYLDRLNLEKN 420
DB 367 SILSILIEVSPPELVCSMLHCSGTRLPALTTHVTOPKDGCFCEVCKKLGVYLDRLNLEKN 426
QY 421 STKOEIILAALKEKGSFLPDPYOKOCQDFVAEYEPVLEILVEYMDSPFVCLKTGACPSAH 480
DB 427 STKOEIILAALKEKGSFLPDPYOKOCQDFVAEYEPVLEILVEYMDSPFVCLKTGACPSAH 486
QY 481 KPLIGTEKCIWGPBYWCQNTETAAQCNAVEHCKRRHVN 518
DB 487 KPLIGTEKCIWGPBYWCQNTETAAQCNAVEHCKRRHVN 524
|||||

Search completed: May 5, 2004, 13:31:04
Job time : 73.3134 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:10:37 ; Search time 22.6287 Seconds
(without alignments)
1181.786 Million cell updates/sec

Title: US-09-743-684a-1_COPY_7_524

Perfect score: 2759

Sequence: 1 LASLIGALAGPVGLKECT.....NTETRAQCNVHCKRHWN 518

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6C_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/6D_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2759	100.0	524	US-09-352-548-1	Sequence 1, Appli
2	2736.5	99.2	523	US-08-100-247-2	Sequence 2, Appli
3	2736.5	99.2	523	US-08-756-031-2	Sequence 2, Appli
4	2729.5	98.9	523	US-08-232-513A-3	Sequence 3, Appli
5	2727.5	98.9	523	US-08-483-146A-2	Sequence 2, Appli
6	2727.5	98.9	523	US-08-484-594A-2	Sequence 2, Appli
7	2727.5	98.9	523	US-09-076-258A-2	Sequence 2, Appli
8	431	15.6	81	US-09-352-548-2	Sequence 2, Appli
9	418.5	15.2	80	US-08-584-671-15	Sequence 15, Appli
10	418.5	15.2	80	US-09-027-376-15	Sequence 15, Appli
11	415.5	15.1	80	US-09-094-192-15	Sequence 15, Appli
12	412	14.9	80	US-08-100-247-3	Sequence 3, Appli
13	412	14.9	80	US-08-483-146A-3	Sequence 3, Appli
14	412	14.9	80	US-08-232-513A-4	Sequence 4, Appli
15	412	14.9	80	US-08-484-594A-3	Sequence 3, Appli
16	412	14.9	80	US-09-076-258A-3	Sequence 3, Appli
17	412	14.9	80	US-08-756-031-3	Sequence 3, Appli
18	341.5	12.4	381	US-09-193-877-2	Sequence 3, Appli
19	327	11.9	61	US-08-584-671-13	Sequence 13, Appli
20	327	11.9	61	US-09-027-376-13	Sequence 13, Appli
21	327	11.9	61	US-09-027-376-13	Sequence 13, Appli
22	327	11.9	61	US-09-027-376-13	Sequence 13, Appli
23	327	11.9	61	US-09-027-376-13	Sequence 13, Appli
24	322	11.7	79	US-08-584-671-16	Sequence 16, Appli
25	322	11.7	79	US-09-027-376-16	Sequence 16, Appli
26	321	11.6	79	US-09-094-192-16	Sequence 16, Appli
27	321	11.6	79	US-08-584-671-14	Sequence 14, Appli

28	321	11.6	79	3	US-09-094-192-14	Sequence 14, Appli
29	268.5	9.7	257	2	US-08-596-684F-7	Sequence 7, Appli
30	239.5	8.7	60	2	US-08-584-671-12	Sequence 12, Appli
31	239.5	8.7	60	3	US-09-027-376-12	Sequence 12, Appli
32	239.5	8.7	60	3	US-09-268-070-4	Sequence 4, Appli
33	239.5	8.7	60	3	US-09-094-192-12	Sequence 12, Appli
34	238.5	8.6	68	3	US-09-268-070-1	Sequence 1, Appli
35	208	7.5	67	3	US-09-268-070-3	Sequence 3, Appli
36	168.5	6.1	181	3	US-08-848-580-12	Sequence 12, Appli
37	168.5	6.1	181	4	US-08-488-123-12	Sequence 12, Appli
38	126.5	4.6	514	4	US-09-107-532A-6282	Sequence 6282, Ap
39	115.5	4.2	1002	4	US-09-762-724-4	Sequence 4, Appli
40	113	4.1	1027	4	US-09-762-724-8	Sequence 8, Appli
41	113	4.1	1027	3	US-08-973-462-8	Sequence 8, Appli
42	112.5	4.1	1014	4	US-09-762-724-2	Sequence 2, Appli
43	111	4.0	22	1	US-08-100-247-1	Sequence 1, Appli
44	111	4.0	22	1	US-08-483-146A-1	Sequence 1, Appli
45	111	4.0	22	1	US-08-483-146A-11	Sequence 11, Appli

ALIGNMENTS

```
RESULT 1
US-09-352-548-1
; Sequence 1, Application US/09352548
; Patent No. 6500431
; GENERAL INFORMATION:
; APPLICANT: G111, Parkash S.
; APPLICANT: Parkash S. G111, M.D., Inc.
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
; FILE REFERENCE: 017986-000410US
; CURRENT APPLICATION NUMBER: US/09/352,548
; EARLIER FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: US 60/092,647
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: prosaposin
; NAME/KEY: PEPTIDE
; LOCATION: (195)..(275)
; OTHER INFORMATION: Saposin B
US-09-352-548-1
Query Match 100.0%; Score 2759; DB 4; Length 524;
Best Local Similarity 100.0%; Pred. No. 9.86-240;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 LASLIGALAGPVGLKECTGSAWCVNTASDCGAKHCLQVWKKPTKSLPCDIC 60
7 LASLIGALAGPVGLKECTGSAWCVNTASDCGAKHCLQVWKKPTKSLPCDIC 66
61 KDVTAAADMLKDNATEEELIVYLEKTCMDLPRKPMASCKEIVDSYLPVILDIITKGENS 120
67 KDVTAAADMLKDNATEEELIVYLEKTCMDLPRKPMASCKEIVDSYLPVILDIITKGENS 126
121 RPEVCASALNLCESLQKHIAELNHQKLESKKIPELDMTEVAVAPPMANIPLLLYPQDGR 180
127 RPEVCASALNLCESLQKHIAELNHQKLESKKIPELDMTEVAVAPPMANIPLLLYPQDGR 186
181 SKPOPKNDGVCDICQWVTDIOTAVRTNSFVQALVHVHVEBCDRLPGMADICNYIS 240
187 SKPOPKNDGVCDICQWVTDIOTAVRTNSFVQALVHVHVEBCDRLPGMADICNYIS 246
241 QYSEIATIGMMHMQPKETCALVGFCEVKEKPMQTLVPAKVASKNVITPALEIVERTKGE 300
247 QYSEIATIGMMHMQPKETCALVGFCEVKEKPMQTLVPAKVASKNVITPALEIVERTKGE 306
```

QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDADPKMSKLPKSLSECCQEVVDYGS 360
DB 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDADPKMSKLPKSLSECCQEVVDYGS 366
QY 361 SIISILILEEVSPELVCSMLHLCGSTRLPALTVAHTQPKDGFCEVCKLVGYIDRLNLEKN 420
DB 367 SIISILILEEVSPELVCSMLHLCGSTRLPALTVAHTQPKDGFCEVCKLVGYIDRLNLEKN 426
QY 421 STQOEIIAALAEKGCSTLPDPYQKCDQFAVEYEPVLIEILVEVMDPSFVCLKIGACPSAH 480
DB 427 STQOEIIAALAEKGCSTLPDPYQKCDQFAVEYEPVLIEILVEVMDPSFVCLKIGACPSAH 486
QY 481 KPILGTEKCIWGPSTYCONTEETAOCNAVEHCKRGHWN 518
DB 487 KPILGTEKCIWGPSTYCONTEETAOCNAVEHCKRGHWN 524

RESULT 2
US-08-100-247-2
Sequence 2, Application US/08100247
Patent No. 5571787
GENERAL INFORMATION:
APPLICANT: O'BRIEN, JOHN S.
APPLICANT: KISHIMOTO, YASUO
TITLE OF INVENTION: PROSAPOSIN AS A NEUTROTROPIC FACTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/100,247
FILING DATE: 19930730
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: O'Brien.002A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
IMMEDIATE SOURCE:
CLONE: PROSAPOSIN
US-08-100-247-2

Query Match 99.2%; Score 2736.5; DB 1; Length 523;
Best Local Similarity 99.4%; Pred. No. 1e-237;
Matches 515; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 LASILGALAGPYLGLKECTRGSAVWCONKTAASDCAVHCICQTMNKPVTYSLPCDTC 60
DB 7 LASILGALAGPYLGLKECTRGSAVWCONKTAASDCAVHCICQTMNKPVTYSLPCDTC 66
QY 61 KDVVTAAGDMLKDNATEEILVLYLEKTCDWLPRNMSASCKEIVDSVLPVILDIIGEMS 120

DB 67 KDVVTAAGDMLKDNATEEILVLYLEKTCDWLPRNMSASCKEIVDSVLPVILDIIGEMS 126
QY 121 RRGVCSALNLCESIQHIAELNHQKQLESNKIPDELMTVEVAPFMANIPLLLYPDQDPR 180
DB 127 RRGVCSALNLCESIQHIAELNHQKQLESNKIPDELMTVEVAPFMANIPLLLYPDQDPR 186
QY 181 SKPQKNDGVCODCICQWVDIOTAVRTNSFVQALVEHKEEEDRLGRGMADICRYIIS 240
DB 187 SKPQKNDGVCODCICQWVDIOTAVRTNSFVQALVEHKEEEDRLGRGMADICRYIIS 245
QY 241 QYSEIAIQMMHMQPKXICALVGCDEKEMPMQTLVPAKVASKNVLPALVELPIKHE 300
DB 246 QYSEIAIQMMHMQPKXICALVGCDEKEMPMQTLVPAKVASKNVLPALVELPIKHE 305
QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDADPKMSKLPKSLSECCQEVVDYGS 360
DB 306 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDADPKMSKLPKSLSECCQEVVDYGS 365
QY 361 SIISILILEEVSPELVCSMLHLCGSTRLPALTVAHTQPKDGFCEVCKLVGYIDRLNLEKN 420
DB 366 SIISILILEEVSPELVCSMLHLCGSTRLPALTVAHTQPKDGFCEVCKLVGYIDRLNLEKN 425
QY 421 STQOEIIAALAEKGCSTLPDPYQKCDQFAVEYEPVLIEILVEVMDPSFVCLKIGACPSAH 480
DB 426 STQOEIIAALAEKGCSTLPDPYQKCDQFAVEYEPVLIEILVEVMDPSFVCLKIGACPSAH 485
QY 481 KPILGTEKCIWGPSTYCONTEETAOCNAVEHCKRGHWN 518
DB 486 KPILGTEKCIWGPSTYCONTEETAOCNAVEHCKRGHWN 523

RESULT 3
US-08-756-031-2
Sequence 2, Application US/08756031
Patent No. 6590074
GENERAL INFORMATION:
APPLICANT: O'BRIEN, JOHN S.
APPLICANT: KISHIMOTO, YASUO
TITLE OF INVENTION: PROSAPOSIN AS A NEUTROTROPIC FACTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/756,031
FILING DATE: 26-NOV-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/100,247
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: O'Brien.002A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
IMMEDIATE SOURCE:
CLONE: PROSAPOSIN
US-08-756-031-2

Query Match 99.2%; Score 2736.5; DB 4; Length 523;
Best Local Similarity 99.4%; Pred. No. 1e-237;
Matches 515; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 LASLIGALAGPVGLKECTRGSAWCONVTASDCGAVKICLOTWNNKPTVKSIPCDIC 60
DB 7 LASLIGALAGPVGLKECTRGSAWCONVTASDCGAVKICLOTWNNKPTVKSIPCDIC 66
QY 61 KDVTAAAGDMKDNATEEELIVLEKTCMDLPRKPMASCKEIVDSYLPVLDIIGKEMS 120
DB 67 KDVTAAAGDMKDNATEEELIVLEKTCMDLPRKPMASCKEIVDSYLPVLDIIGKEMS 126
QY 121 RRGVCSALNLCESIQKHLAELNHQKLESNKIPELDMTEVVAAPPMANIPLLVPODGP 180
DB 127 RRGVCSALNLCESIQKHLAELNHQKLESNKIPELDMTEVVAAPPMANIPLLVPODGP 186
QY 181 SKPQKNDGVCQDCIQWTDIQTAVRTNSTFVQALVEHVKEBCDRLPGMADICKNYIS 240
DB 187 SKPQKNDGVCQDCIQWTDIQTAVRTNSTFVQALVEHVKEBCDRLPGMADICKNYIS 245
QY 241 QYSEIAIOMMMHMPKEICALVGFCEVKEPMQTLVPAKASKNVLPALDLVPIKXHE 300
DB 246 QYSEIAIOMMMHMPKEICALVGFCEVKEPMQTLVPAKASKNVLPALDLVPIKXHE 305
QY 301 VPAKSDVYCEVCEPLVKEVTKLIDNNKTEKEILDAFDMCKSLPKSLSEECQEVVDITYGS 360
DB 306 VPAKSDVYCEVCEPLVKEVTKLIDNNKTEKEILDAFDMCKSLPKSLSEECQEVVDITYGS 365
QY 361 SILSILAEVSPBELVCSMLHLCSTGRPLPALTVHTVQPKDGFCEVCKKLVGLDRLNLEKN 420
DB 366 SILSILAEVSPBELVCSMLHLCSTGRPLPALTVHTVQPKDGFCEVCKKLVGLDRLNLEKN 425
QY 421 STKOEIIAALEKGSFLPDYQKQCDQFVAEYEPVLLEIIVWDPSPVCLIKIGACPSAH 480
DB 426 STKOEIIAALEKGSFLPDYQKQCDQFVAEYEPVLLEIIVWDPSPVCLIKIGACPSAH 485
QY 481 KPLIGTEKCIWGPSYWCNTETAACNAVEHCKRHVNN 518
DB 486 KPLIGTEKCIWGPSYWCNTETAACNAVEHCKRHVNN 523

RESULT 4

US-08-232-513A-3
Sequence 3, Application US/08232513A
Patent No. 5700909
GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides
TITLE OF INVENTION: as Therapeutic Agents
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232, 513A
FILING DATE: 21-APR-1994

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/100,247
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-UD 1643
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..523
OTHER INFORMATION: /label= Hum_prosaposin
US-08-232-513A-3

Query Match 98.9%; Score 2729.5; DB 1; Length 523;
Best Local Similarity 99.2%; Pred. No. 4.4e-237;
Matches 514; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 LASLIGALAGPVGLKECTRGSAWCONVTASDCGAVKICLOTWNNKPTVKSIPCDIC 60
DB 7 LASLIGALAGPVGLKECTRGSAWCONVTASDCGAVKICLOTWNNKPTVKSIPCDIC 66
QY 61 KDVTAAAGDMKDNATEEELIVLEKTCMDLPRKPMASCKEIVDSYLPVLDIIGKEMS 120
DB 67 KDVTAAAGDMKDNATEEELIVLEKTCMDLPRKPMASCKEIVDSYLPVLDIIGKEMS 126
QY 121 RRGVCSALNLCESIQKHLAELNHQKLESNKIPELDMTEVVAAPPMANIPLLVPODGP 180
DB 127 RRGVCSALNLCESIQKHLAELNHQKLESNKIPELDMTEVVAAPPMANIPLLVPODGP 186
QY 181 SKPQKNDGVCQDCIQWTDIQTAVRTNSTFVQALVEHVKEBCDRLPGMADICKNYIS 240
DB 187 SKPQKNDGVCQDCIQWTDIQTAVRTNSTFVQALVEHVKEBCDRLPGMADICKNYIS 245
QY 241 QYSEIAIOMMMHMPKEICALVGFCEVKEPMQTLVPAKASKNVLPALDLVPIKXHE 300
DB 246 QYSEIAIOMMMHMPKEICALVGFCEVKEPMQTLVPAKASKNVLPALDLVPIKXHE 305
QY 301 VPAKSDVYCEVCEPLVKEVTKLIDNNKTEKEILDAFDMCKSLPKSLSEECQEVVDITYGS 360
DB 306 VPAKSDVYCEVCEPLVKEVTKLIDNNKTEKEILDAFDMCKSLPKSLSEECQEVVDITYGS 365
QY 361 SILSILAEVSPBELVCSMLHLCSTGRPLPALTVHTVQPKDGFCEVCKKLVGLDRLNLEKN 420
DB 366 SILSILAEVSPBELVCSMLHLCSTGRPLPALTVHTVQPKDGFCEVCKKLVGLDRLNLEKN 425
QY 421 STKOEIIAALEKGSFLPDYQKQCDQFVAEYEPVLLEIIVWDPSPVCLIKIGACPSAH 480
DB 426 STKOEIIAALEKGSFLPDYQKQCDQFVAEYEPVLLEIIVWDPSPVCLIKIGACPSAH 485
QY 481 KPLIGTEKCIWGPSYWCNTETAACNAVEHCKRHVNN 518
DB 486 KPLIGTEKCIWGPSYWCNTETAACNAVEHCKRHVNN 523

RESULT 5

US-08-483-146A-2
Sequence 2, Application US/08483146A
Patent No. 5696080
GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED

TITLE OF INVENTION: THEREFROM
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Blvd. 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,146A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.002DV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-483-146A-2

Query Match 98.9%; Score 2727.5; DB 1; Length 523;
Best Local Similarity 99.2%; Pred. No. 6.7e-237;
Matches 514; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
QY 1 LASTLGALLAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLOTWNKPTVKSLPEDIC 60
DB 7 LASTLGALLAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLOTWNKPTVKSLPEDIC 66
QY 61 KDVTAAAGDMKDNATEEELIVYEKTCMDLPKPNMSASCKEIVDSYLPVILDIITKEMS 120
DB 67 KDVTAAAGDMKDNATEEELIVYEKTCMDLPKPNMSASCKEIVDSYLPVILDIITKEMS 126
QY 121 RRGVCSALNLCESIQKHLAEINHQKQLESNKIPELDMTEVVAFPMANIPLLIYPQDGP 180
DB 127 RRGVCSALNLCESIQKHLAEINHQKQLESNKIPELDMTEVVAFPMANIPLLIYPQDGP 186
QY 181 SKPPKNGVGDCCICQMTDIQTAVRTNSTFVQALVEHVEKCDRLGPGMADICKNYIS 240
DB 187 SKPPKNGVGDCCICQMTDIQTAVRTNSTFVQALVEHVEKCDRLGPGMADICKNYIS 245
QY 241 QYSIAIQMMHMQPKSICALVGFCDVEMKPMQTLVPAVASKNYIPALDVLDPKKEH 300
DB 246 QYSIAIQMMHMQPKSICALVGFCDVEMKPMQTLVPAVASKNYIPALDVLDPKKEH 305
QY 301 VPAASDVYCEVCELVKEVTKLIDNNKTEKEIIDAFPKMCKSLPKSLSECCQEVVDYGS 360
DB 306 VPAASDVYCEVCELVKEVTKLIDNNKTEKEIIDAFPKMCKSLPKSLSECCQEVVDYGS 365
QY 361 SIISILIEVSPELVCSMLHLCSTRLPALTVHTOPKOGFCEVCKKIVGYIDRLKLN 420
DB 366 SIISILIEVSPELVCSMLHLCSTRLPALTVHTOPKOGFCEVCKKIVGYIDRLKLN 425
QY 421 STGOEIIAALBKGCSPDPYQKQCFVAEYEPVLEILVEVWDSPFVCLKIGACPSAH 480

DB 426 STGOEIIAALBKGCSPDPYQKQCFVAEYEPVLEILVEVWDSPFVCLKIGACPSAH 485
QY 481 KPLGTEKICWGPWYQONTTAAQCNVAHECKRHVN 518
DB 486 KPLGTEKICWGPWYQONTTAAQCNVAHECKRHVN 523
RESULT 6
US-08-484-594A-2
Sequence 2, Application US/08484594A
Patent No. 5714459
GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
APPLICANT: Kishimoto, Yasuo
TITLE OF INVENTION: USE OF PROSAPIN AND NEUROTROPHIC PEPTIDES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,594A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/100,247
FILING DATE: 30-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS.002DV2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-484-594A-2
Query Match 98.9%; Score 2727.5; DB 1; Length 523;
Best Local Similarity 99.2%; Pred. No. 6.7e-237;
Matches 514; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
QY 1 LASTLGALLAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLOTWNKPTVKSLPEDIC 60
DB 7 LASTLGALLAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLOTWNKPTVKSLPEDIC 66
QY 61 KDVTAAAGDMKDNATEEELIVYEKTCMDLPKPNMSASCKEIVDSYLPVILDIITKEMS 120
DB 67 KDVTAAAGDMKDNATEEELIVYEKTCMDLPKPNMSASCKEIVDSYLPVILDIITKEMS 126
QY 121 RRGVCSALNLCESIQKHLAEINHQKQLESNKIPELDMTEVVAFPMANIPLLIYPQDGP 180
DB 127 RRGVCSALNLCESIQKHLAEINHQKQLESNKIPELDMTEVVAFPMANIPLLIYPQDGP 186
QY 181 SKPPKNGVGDCCICQMTDIQTAVRTNSTFVQALVEHVEKCDRLGPGMADICKNYIS 240
DB 187 SKPPKNGVGDCCICQMTDIQTAVRTNSTFVQALVEHVEKCDRLGPGMADICKNYIS 245

QY 241 QYSEIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIALPDLVEPIKKE 300
DB 246 QYSEIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIALPDLVEPIKKE 305
QY 301 VPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDMKCSKLPKSLSEECOEVDITYGS 360
DB 306 VPASDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDMKCSKLPKSLSEECOEVDITYGS 365
QY 361 SILSILLESVPELVCSMLHLCSTRLPALIVHTVTPKDGFCFCEVCKLVGYLDRLNEKN 420
DB 366 SILSILLESVPELVCSMLHLCSTRLPALIVHTVTPKDGFCFCEVCKLVGYLDRLNEKN 425
QY 421 STKOEIIAALKEKGSFLPDYQKCDQFVAEYEPVLIIEIVWMPSPVCLIKIGACPSAH 480
DB 426 STKOEIIAALKEKGSFLPDYQKCDQFVAEYEPVLIIEIVWMPSPVCLIKIGACPSAH 485
QY 481 KPLIGTEKICMGPSYWCNTETAACNAVEHCKRHVN 518
DB 486 KPLIGTEKICMGPSYWCNTETAACNAVEHCKRHVN 523

RESULT 7
US-09-743-684a-2
Sequence 2, Application US/09076258A

Patent No. 6559124

GENERAL INFORMATION:

APPLICANT: O'Brien, John S.

APPLICANT: Kishimoto, Yasuo

TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS

TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbels, Martens, Olson and Bear

STREET: 620 Newport Center Blvd. 16th Floor

CITY: Newport Beach

STATE: CA

COUNTRY: USA

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/076,258A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/958,970

FILING DATE: 28-OCT-97

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/483,146

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: 08/100,247

FILING DATE: 30-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: Israelien, Ned A

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: MYELOS.2DVIC2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-235-8550

TELEFAX: 619-235-0176

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 523 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: N-terminal

US-09-743-684a-2

Query Match

Best Local Similarity 98.94; Score 2727.5; DB 4; Length 523;

Matches 514; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 LASLIGALAGPVLGKCTRGSAVWCNVNTASDCGVRKCLQTVNKPVTYSLPCIC 60
DB 7 LASLIGALAGPVLGKCTRGSAVWCNVNTASDCGVRKCLQTVNKPVTYSLPCIC 66
QY 61 KDVTYAGDMKDNATEBEILVYLEKTCMDLPKRNMSASCKEIVDSYLPVLLIIKGRMS 120
DB 67 KDVTYAGDMKDNATEBEILVYLEKTCMDLPKRNMSASCKEIVDSYLPVLLIIKGRMS 126
QY 121 RPEVCSALNLCESLQKHLAELNHQKLESKIKIIELDTEVVAAPPMANIPILLIPQDPR 180
DB 127 RPEVCSALNLCESLQKHLAELNHQKLESKIKIIELDTEVVAAPPMANIPILLIPQDPR 186
QY 181 SKPQKNGDVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADICKNYIS 240
DB 187 SKPQKNGDVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADICKNYIS 245
QY 241 QYSEIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIALPDLVEPIKKE 300
DB 246 QYSEIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIALPDLVEPIKKE 305
QY 301 VPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDMKCSKLPKSLSEECOEVDITYGS 360
DB 306 VPASDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDMKCSKLPKSLSEECOEVDITYGS 365
QY 361 SILSILLESVPELVCSMLHLCSTRLPALIVHTVTPKDGFCFCEVCKLVGYLDRLNEKN 420
DB 366 SILSILLESVPELVCSMLHLCSTRLPALIVHTVTPKDGFCFCEVCKLVGYLDRLNEKN 425
QY 421 STKOEIIAALKEKGSFLPDYQKCDQFVAEYEPVLIIEIVWMPSPVCLIKIGACPSAH 480
DB 426 STKOEIIAALKEKGSFLPDYQKCDQFVAEYEPVLIIEIVWMPSPVCLIKIGACPSAH 485
QY 481 KPLIGTEKICMGPSYWCNTETAACNAVEHCKRHVN 518
DB 486 KPLIGTEKICMGPSYWCNTETAACNAVEHCKRHVN 523

RESULT 8

US-09-352-548-2

Sequence 2, Application US/09352548

Patent No. 6500431

GENERAL INFORMATION:

APPLICANT: Gill, Parkash S.

APPLICANT: Parkash S. Gill, M.D., Inc.

TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth

FILE REFERENCE: 017986-000410US

CURRENT APPLICATION NUMBER: US/09/352,548

CURRENT FILING DATE: 1999-07-12

EARLIER APPLICATION NUMBER: US 60/092,647

EARLIER FILING DATE: 1998-07-13

NUMBER OF SEQ ID NOS: 59

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 81

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

INFORMATION: Saposin B

US-09-352-548-2

Query Match 15.64; Score 431; DB 4; Length 81;

Best Local Similarity 100.0%; Pred. No. 1.7e-31;

Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 GDVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADICKNYISQYSEIAIQ 248
DB 1 GDVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADICKNYISQYSEIAIQ 60

QY 249 MAMHQPKEICALVGFCEYK 269
Db 61 MAMHQPKEICALVGFCEYK 81

RESULT 9

US-08-584-671-15
Sequence 15, Application US/08584671
Patent No. 5910568
GENERAL INFORMATION:
APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM
TITLE OF INVENTION: TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA
ADDRESS: STATE UNIVERSITY
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/584,671
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MONAHAN, THOMAS J
REGISTRATION NUMBER: 29835
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 814-865-6277
TELEFAX: 814-865-3591
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 80
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: UNKNOWN
US-08-584-671-15

Query Match 15.2%; Score 418.5; DB 2; Length 80;
Best Local Similarity 98.8%; Pred. No. 2.3e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 122 PGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAPFMANIPLLLYPDGPRS 181
Db 1 PGEVC-ALNLCESLQKHLAELNHQKLESNKIPELDMTEVAPFMANIPLLLYPDGPRS 59
QY 182 KPQPKNDGVCDICIQMTDI 202
Db 60 KPQPKNDGVCDICIQMTDI 80

RESULT 10

US-09-027-376-15
Sequence 15, Application US/09027376
Patent No. 6004586
GENERAL INFORMATION:
APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM
TITLE OF INVENTION: TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE
NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA
ADDRESS: STATE UNIVERSITY
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/027,376
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/584,671
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MONAHAN, THOMAS J
REGISTRATION NUMBER: 29835
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 814-865-6277
TELEFAX: 814-865-3591
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 80
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: UNKNOWN
US-09-027-376-15

Query Match 15.2%; Score 418.5; DB 3; Length 80;
Best Local Similarity 98.8%; Pred. No. 2.3e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 122 PGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAPFMANIPLLLYPDGPRS 181
Db 1 PGEVC-ALNLCESLQKHLAELNHQKLESNKIPELDMTEVAPFMANIPLLLYPDGPRS 59
QY 182 KPQPKNDGVCDICIQMTDI 202
Db 60 KPQPKNDGVCDICIQMTDI 80

RESULT 11

US-09-094-192-15
Sequence 15, Application US/09094192
Patent No. 6103483
GENERAL INFORMATION:
APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM TO EGG SURFACES AND PROCI
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA STATE UNIVERSITY
STREET: 113 TECHNOLOGY CENTER
CITY: UNIVERSITY PARK
STATE: PENNSYLVANIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 16802-7000
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: NEC 286
OPERATING SYSTEM: DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/094,192
FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: MONAHAN, THOMAS J
REGISTRATION NUMBER: 29835
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 814-865-6277
TELEFAX: 814-865-3591
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 80
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: UNKNOWN
US-09-094-192-15

Query Match 15.1%; Score 415.5; DB 3; Length 80;
Best Local Similarity 97.5%; Pred. No. 4.3e-30;
Matches 79; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 122 PGEVCSALNLCESLQKHLAEINHQKLSNKKIPELDMTEVVAPEMANIPILLYPODGPBS 181
Db 1 PGEVCSALNLCESLQKHLAEINHQKLSNKKIPELDMTEVVAPEMANIPILLYPODGPBS 59
Qy 182 KPQKNGDVCCDCIQWTDI 202
Db 60 KPQKNGDVCCDCIQWTDI 80

RESULT 12
US-08-100-247-3
Sequence 3, Application US/08100247
Patent No. 5571787

GENERAL INFORMATION:
APPLICANT: O'BRIEN, JOHN S.
APPLICANT: KISHIMOTO, YASUO
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: KNOBBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/100,247
FILING DATE: 19930730

CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: O'BRIEN, 002A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
IMMEDIATE SOURCE:
CLONE: SAPOSIN C
US-08-100-247-3

Query Match 14.9%; Score 412; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.8e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 SDVYCEVEFLVKEVTKLIDNNKTEKEILDAFDKXCSLPSLSSECCQVVDYTGSSILS 364
Db 1 SDVYCEVEFLVKEVTKLIDNNKTEKEILDAFDKXCSLPSLSSECCQVVDYTGSSILS 60

Qy 365 ILLEEVSPELVCSMLHLCSG 384
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 13
US-08-483-146A-3
Sequence 3, Application US/08483146A
Patent No. 5696080

GENERAL INFORMATION:
APPLICANT: O'Brien, John S.
APPLICANT: Kishimoto, Yasuo
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Blvd. 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,146A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: MYELOS, 002DV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-483-146A-3

Query Match 14.9%; Score 412; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.8e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 SDVYCEVEFLVKEVTKLIDNNKTEKEILDAFDKXCSLPSLSSECCQVVDYTGSSILS 364
Db 1 SDVYCEVEFLVKEVTKLIDNNKTEKEILDAFDKXCSLPSLSSECCQVVDYTGSSILS 60

Qy 365 ILLEEVSPELVCSMLHLCSG 384
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 14
US-08-232-513A-4
; Sequence 4, Application US/08232513A
; Patent No. 5700909
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; TITLE OF INVENTION: Proasoposin and Cytokine-Derived Peptides
; TITLE OF INVENTION: as Therapeutic Agents
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,513A
; FILING DATE: 21-APR-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/100,247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 1643
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-8949
; TELEFAX: (619) 535-9001
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..80
; OTHER INFORMATION: /label= Saposin_C
US-08-232-513A-4
Query Match 14.9%; Score 412; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.8e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 305 SDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDADFDMKSKLPKSLSECEVVDYSSILS 364
DB 1 SDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDADFDMKSKLPKSLSECEVVDYSSILS 60
QY 365 ILLEEVSPELVCSMLHLCSG 384
DB 61 ILLEEVSPELVCSMLHLCSG 80
RESULT 15
US-08-484-594A-3
; Sequence 3, Application US/08484594A
; Patent No. 5714459
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; APPLICANT: Kishimoto, Yasuo
; TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUROTROPHIC PEPTIDES
; TITLE OF INVENTION: DERIVED THEREFROM
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:

ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,594A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/100,247
; FILING DATE: 30-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelien, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: MYELOS.002DV2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
US-08-484-594A-3
Query Match 14.9%; Score 412; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 8.8e-30;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 305 SDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDADFDMKSKLPKSLSECEVVDYSSILS 364
DB 1 SDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDADFDMKSKLPKSLSECEVVDYSSILS 60
QY 365 ILLEEVSPELVCSMLHLCSG 384
DB 61 ILLEEVSPELVCSMLHLCSG 80

Search completed: May 5, 2004, 13:16:44
Job time : 23.6287 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:14:53 ; Search time 55.0959 Seconds

(without alignments)
2606.077 Million cell updates/sec

Title: US-09-743-684a-1_COPY_7_524

Perfect score: 2759
Sequence: 1 LASLIGALAPVIGLKECT.....NTETAAQCNAYEHCHRYWN 518

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 27189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09C_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2759	100.0	524	9	US-09-870-759-60
2	2759	100.0	524	10	US-09-751-708A-60
3	2759	100.0	524	12	US-10-267-502-386
4	2747.5	99.6	527	9	US-09-870-759-61
5	2747.5	99.6	527	10	US-09-751-708A-61
6	2747.5	99.6	527	14	US-10-060-036-73
7	2727.5	98.9	523	9	US-09-767-007A-2
8	2419.5	87.7	479	10	US-09-978-418-40
9	1947	70.6	554	14	US-10-205-194-176
10	1189	43.1	521	15	US-10-276-162-1
11	1127.5	40.9	531	11	US-09-833-245-903
12	1093	38.6	209	14	US-10-043-467-340
13	822.5	29.8	362	12	US-10-332-426-8
14	549.5	19.9	953	12	US-10-267-502-385
15	545	19.8	241	15	US-10-108-260A-4529

16	539.5	19.6	156	12	US-09-925-298-644	Sequence 644, App
17	539.5	19.6	156	14	US-10-102-806-644	Sequence 644, App
18	412	14.9	80	9	US-09-767-007A-3	Sequence 3, Appl1
19	412	14.9	80	9	US-09-753-126-3	Sequence 3, Appl1
20	412	14.9	80	15	US-10-330-697-3	Sequence 3, Appl1
21	412	14.9	592	9	US-09-753-126-4	Sequence 4, Appl1
22	412	14.9	592	15	US-10-330-697-4	Sequence 4, Appl1
23	340.5	12.3	381	15	US-10-236-031B-10	Sequence 10, Appl1
24	340.5	12.3	381	15	US-10-295-027-1223	Sequence 1223, Ap
25	235.5	8.3	216	12	US-10-424-599-276328	Sequence 276328
26	211	7.6	40	10	US-09-780-438A-1	Sequence 1, Appl1
27	211	7.6	212	12	US-10-424-599-157904	Sequence 157904, A
28	204.5	7.4	246	12	US-10-425-114-41077	Sequence 41077, A
29	201	7.3	38	10	US-09-780-438A-2	Sequence 2, Appl1
30	201	7.3	212	12	US-10-424-599-157906	Sequence 157906, A
31	197	7.1	252	12	US-10-425-114-55606	Sequence 55606, A
32	184	6.7	243	12	US-10-425-114-67929	Sequence 67929, A
33	175.5	6.4	265	12	US-10-425-114-51535	Sequence 51535, A
34	169.5	6.1	236	6	US-10-424-599-227975	Sequence 227975, A
35	168.5	6.1	181	8	US-08-488-123-12	Sequence 12, Appl1
36	165.5	6.0	246	12	US-10-425-114-60351	Sequence 60351, A
37	163	5.9	514	14	US-10-339-351-3	Sequence 3, Appl1
38	162	5.9	514	14	US-10-339-351-1	Sequence 1, Appl1
39	159.5	5.8	281	12	US-10-424-599-205602	Sequence 205602, A
40	157.5	5.7	282	12	US-10-424-599-205603	Sequence 205603, A
41	156.5	5.7	507	12	US-10-424-599-245585	Sequence 245585, A
42	152	5.5	471	12	US-10-425-114-71447	Sequence 71427, A
43	151	5.5	393	12	US-10-425-114-65142	Sequence 65142, A
44	147.5	5.3	223	12	US-10-425-114-40454	Sequence 40454, A
45	146.5	5.3	295	12	US-10-425-114-57134	Sequence 57134, A

ALIGNMENTS

RESULT 1
US-09-870-759-60
Sequence 60, Application US/09870759
Patent No. US20020177551A1
GENERAL INFORMATION:
APPLICANT: TERMAN, David S
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
FILE REFERENCE: 870759
CURRENT FILING DATE: 2002-01-14
PRIOR APPLICATION NUMBER: US 60/208,128
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 166
SOFTWARE: PatentIn version 3.1
SEQ ID NO 60
LENGTH: 524
TYPE: PRT
ORGANISM: Homo sapiens
US-09-870-759-60

Query Match 100.0% Score 2759; DB 9; Length 524;
Best Local Similarity 100.0%; Pred. No. 5.1e-237;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LASLIGALAPVIGLKECTGSAVWCONVTASDCGAVKHCLQTVWKKPTVKSJPCIC	60
DB	7	LASLIGALAPVIGLKECTGSAVWCONVTASDCGAVKHCLQTVWKKPTVKSJPCIC	66
QY	61	KDVVTAAGDMKDNATEEEIIIVLEKTCMDLKPKNMSASCKEIVDSYLPVILDIKGBMS	120
DB	67	KDVVTAAGDMKDNATEEEIIIVLEKTCMDLKPKNMSASCKEIVDSYLPVILDIKGBMS	126
QY	121	RPEVCALNLCESIQKLAELNOKOESKIKELDMTEVAPPMANIPILLYPODSBR	180
DB	127	RPEVCALNLCESIQKLAELNOKOESKIKELDMTEVAPPMANIPILLYPODSBR	186
QY	181	SKPOPKDGVCDQCICQWTDIQTAVRTNSFVVALVHVHVEBCDRLGPMADICKNYS	240

Db 187 SKPQKNDGVCDQDCIQWVTDIQTAVRTNSTFVQALVEHVEECDRLGPGMADICKNYIS 246
Qy 241 QYSEIATQMMHMQPKKEICALVGFCDVEKEMPMQTLVPARVASKNVIPALETVEPIKKEH 300
Db 247 QYSEIATQMMHMQPKKEICALVGFCDVEKEMPMQTLVPARVASKNVIPALETVEPIKKEH 306
Qy 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIILDAFPMKSKLPKSLSECOEVDVTYGS 360
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIILDAFPMKSKLPKSLSECOEVDVTYGS 366
Qy 361 SILSILIEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGGFCVCKLVGYLDRLNEKN 420
Db 367 SILSILIEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGGFCVCKLVGYLDRLNEKN 426
Qy 421 STKOEIIAALKEKGSFLPDPYQKQDQFVAEYEPVLIEILVEWMDSPFVCLKIGACPSAH 480
Db 427 STKOEIIAALKEKGSFLPDPYQKQDQFVAEYEPVLIEILVEWMDSPFVCLKIGACPSAH 486
Qy 481 KPLIGTEKCIWGPSYWCQNTETAACNAVAHECHKRWAN 518
Db 487 KPLIGTEKCIWGPSYWCQNTETAACNAVAHECHKRWAN 524

RESULT 2

US-09-751-708A-60
; Sequence 60, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-751-708A-60

Query Match 100.0%; Score 2759; DB 10; Length 524;
Best Local Similarity 100.0%; Pred. No. 5,1e-237;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LASILGAALAGPVIGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKSJLPCDIC 60
Db 7 LASILGAALAGPVIGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKSJLPCDIC 66
Qy 61 KDVTAAADMLKDNATEEELIYLEKTCDWLPRKPMNSASCKEIVDSYLPVILDIKSGMS 120
Db 67 KDVTAAADMLKDNATEEELIYLEKTCDWLPRKPMNSASCKEIVDSYLPVILDIKSGMS 126
Qy 121 RPEVCSALNLCESLOKHLAEINHOKEESNKIPBLDMTEVYAPPMANIPILLYPQDGP 180
Db 127 RPEVCSALNLCESLOKHLAEINHOKEESNKIPBLDMTEVYAPPMANIPILLYPQDGP 186
Qy 181 SKPQKNDGVCDQDCIQWVTDIQTAVRTNSTFVQALVEHVEECDRLGPGMADICKNYIS 240
Db 187 SKPQKNDGVCDQDCIQWVTDIQTAVRTNSTFVQALVEHVEECDRLGPGMADICKNYIS 246
Qy 241 QYSEIATQMMHMQPKKEICALVGFCDVEKEMPMQTLVPARVASKNVIPALETVEPIKKEH 300
Db 247 QYSEIATQMMHMQPKKEICALVGFCDVEKEMPMQTLVPARVASKNVIPALETVEPIKKEH 306
Qy 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIILDAFPMKSKLPKSLSECOEVDVTYGS 360
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIILDAFPMKSKLPKSLSECOEVDVTYGS 366
Qy 361 SILSILIEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGGFCVCKLVGYLDRLNEKN 420

Db 367 SILSILIEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGGFCVCKLVGYLDRLNEKN 426
Qy 421 STKOEIIAALKEKGSFLPDPYQKQDQFVAEYEPVLIEILVEWMDSPFVCLKIGACPSAH 480
Db 427 STKOEIIAALKEKGSFLPDPYQKQDQFVAEYEPVLIEILVEWMDSPFVCLKIGACPSAH 486
Qy 481 KPLIGTEKCIWGPSYWCQNTETAACNAVAHECHKRWAN 518
Db 487 KPLIGTEKCIWGPSYWCQNTETAACNAVAHECHKRWAN 524

RESULT 3

US-10-267-502-386
; Sequence 386, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 386
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-267-502-386

Query Match 100.0%; Score 2759; DB 12; Length 524;
Best Local Similarity 100.0%; Pred. No. 5,1e-237;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LASILGAALAGPVIGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKSJLPCDIC 60
Db 7 LASILGAALAGPVIGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKSJLPCDIC 66
Qy 61 KDVTAAADMLKDNATEEELIYLEKTCDWLPRKPMNSASCKEIVDSYLPVILDIKSGMS 120
Db 67 KDVTAAADMLKDNATEEELIYLEKTCDWLPRKPMNSASCKEIVDSYLPVILDIKSGMS 126
Qy 121 RPEVCSALNLCESLOKHLAEINHOKEESNKIPBLDMTEVYAPPMANIPILLYPQDGP 180
Db 127 RPEVCSALNLCESLOKHLAEINHOKEESNKIPBLDMTEVYAPPMANIPILLYPQDGP 186
Qy 181 SKPQKNDGVCDQDCIQWVTDIQTAVRTNSTFVQALVEHVEECDRLGPGMADICKNYIS 240
Db 187 SKPQKNDGVCDQDCIQWVTDIQTAVRTNSTFVQALVEHVEECDRLGPGMADICKNYIS 246
Qy 241 QYSEIATQMMHMQPKKEICALVGFCDVEKEMPMQTLVPARVASKNVIPALETVEPIKKEH 300
Db 247 QYSEIATQMMHMQPKKEICALVGFCDVEKEMPMQTLVPARVASKNVIPALETVEPIKKEH 306
Qy 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIILDAFPMKSKLPKSLSECOEVDVTYGS 360
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIILDAFPMKSKLPKSLSECOEVDVTYGS 366
Qy 361 SILSILIEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGGFCVCKLVGYLDRLNEKN 420
Db 367 SILSILIEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGGFCVCKLVGYLDRLNEKN 426
Qy 421 STKOEIIAALKEKGSFLPDPYQKQDQFVAEYEPVLIEILVEWMDSPFVCLKIGACPSAH 480
Db 427 STKOEIIAALKEKGSFLPDPYQKQDQFVAEYEPVLIEILVEWMDSPFVCLKIGACPSAH 486
Qy 481 KPLIGTEKCIWGPSYWCQNTETAACNAVAHECHKRWAN 518
Db 487 KPLIGTEKCIWGPSYWCQNTETAACNAVAHECHKRWAN 524

RESULT 4

US-09-870-759-61


```
/ Sequence 61, Application US/09870759
/ Patent No. US20020177551A1
/ GENERAL INFORMATION:
/ APPLICANT: TERMAN, David S
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
/ FILE REFERENCE: 870759
/ CURRENT APPLICATION NUMBER: US/09/870,759
/ CURRENT FILING DATE: 2002-01-14
/ PRIOR APPLICATION NUMBER: US 60/208,128
/ PRIOR FILING DATE: 2000-05-30
/ NUMBER OF SEQ ID NOS: 166
/ SOFTWARE: Patent version 3.1
/ SEQ ID NO 61
/ LENGTH: 527
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-870-759-61

Query Match          99.6%; Score 2747.5; DB 9; Length 527;
Best Local Similarity 99.4%; Pred. No. 5.4e-236;
Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 LASLIGALAGPVIGLKECTRGSAVWCQNVKTASDCGAVKHCLOTVWNNKPTVKSIPCDIC 60
DB 7 LASLIGALAGPVIGLKECTRGSAVWCQNVKTASDCGAVKHCLOTVWNNKPTVKSIPCDIC 66
QY 61 KDVTAAAGDMKDNATEEEIIVLYLEKTCDDMLPKPNNMSCKEIVDSYLPVLIIDIIKGMS 120
DB 67 KDVTAAAGDMKDNATEEEIIVLYLEKTCDDMLPKPNNMSCKEIVDSYLPVLIIDIIKGMS 126
QY 121 RPEVCSALNLCESIQKHLAELNHQKQLESNKIPELDMEVAVAPPMANIPILLYPQDGR 180
DB 127 RPEVCSALNLCESIQKHLAELNHQKQLESNKIPELDMEVAVAPPMANIPILLYPQDGR 186
QY 181 SKPQKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 240
DB 187 SKPQKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 246
QY 241 QYSEIAIQMMHMM--OPKEICALVGFCDVEKEMPMQTLVPAKVASKNVITPALEVEPIK 297
DB 247 QYSEIAIQMMHMMQDOQPEKICALVGFCDVEKEMPMQTLVPAKVASKNVITPALEVEPIK 306
QY 298 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSSECOEVDVT 357
DB 307 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSSECOEVDVT 366
QY 358 YGSSILSLILBEVSPBELVCSMLHLCSGTRLPALTVHVTQPKDGGFCEVCKLVGYLDNRL 417
DB 367 YGSSILSLILBEVSPBELVCSMLHLCSGTRLPALTVHVTQPKDGGFCEVCKLVGYLDNRL 426
QY 418 EKNSTKQBIILALEKGCSTFLPDPYQKCDQFAVEYEPVLIETLVWMDPSFVCLKIGACP 477
DB 427 EKNSTKQBIILALEKGCSTFLPDPYQKCDQFAVEYEPVLIETLVWMDPSFVCLKIGACP 486
QY 478 SAHKPLGTEKCIWGPSTYWCNTETAACNAVEHCKRHVMN 518
DB 487 SAHKPLGTEKCIWGPSTYWCNTETAACNAVEHCKRHVMN 527

RESULT 5
US-09-751-708A-61
/ Sequence 61, Application US/09751708A
/ Publication No. US20030157113A1
/ GENERAL INFORMATION:
/ APPLICANT: TERMAN, David S
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
/ FILE REFERENCE: 751708
/ CURRENT APPLICATION NUMBER: US/09/751,708A
/ CURRENT FILING DATE: 2002-10-15
/ PRIOR APPLICATION NUMBER: US 60/173,371
/ PRIOR FILING DATE: 1999-12-28
/ NUMBER OF SEQ ID NOS: 166
/ SOFTWARE: Patent version 3.1
```

```
/ SEQ ID NO 61
/ LENGTH: 527
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-751-708A-61

Query Match          99.6%; Score 2747.5; DB 10; Length 527;
Best Local Similarity 99.4%; Pred. No. 5.4e-236;
Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 LASLIGALAGPVIGLKECTRGSAVWCQNVKTASDCGAVKHCLOTVWNNKPTVKSIPCDIC 60
DB 7 LASLIGALAGPVIGLKECTRGSAVWCQNVKTASDCGAVKHCLOTVWNNKPTVKSIPCDIC 66
QY 61 KDVTAAAGDMKDNATEEEIIVLYLEKTCDDMLPKPNNMSCKEIVDSYLPVLIIDIIKGMS 120
DB 67 KDVTAAAGDMKDNATEEEIIVLYLEKTCDDMLPKPNNMSCKEIVDSYLPVLIIDIIKGMS 126
QY 121 RPEVCSALNLCESIQKHLAELNHQKQLESNKIPELDMEVAVAPPMANIPILLYPQDGR 180
DB 127 RPEVCSALNLCESIQKHLAELNHQKQLESNKIPELDMEVAVAPPMANIPILLYPQDGR 186
QY 181 SKPQKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 240
DB 187 SKPQKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 246
QY 241 QYSEIAIQMMHMM--OPKEICALVGFCDVEKEMPMQTLVPAKVASKNVITPALEVEPIK 297
DB 247 QYSEIAIQMMHMMQDOQPEKICALVGFCDVEKEMPMQTLVPAKVASKNVITPALEVEPIK 306
QY 298 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSSECOEVDVT 357
DB 307 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSSECOEVDVT 366
QY 358 YGSSILSLILBEVSPBELVCSMLHLCSGTRLPALTVHVTQPKDGGFCEVCKLVGYLDNRL 417
DB 367 YGSSILSLILBEVSPBELVCSMLHLCSGTRLPALTVHVTQPKDGGFCEVCKLVGYLDNRL 426
QY 418 EKNSTKQBIILALEKGCSTFLPDPYQKCDQFAVEYEPVLIETLVWMDPSFVCLKIGACP 477
DB 427 EKNSTKQBIILALEKGCSTFLPDPYQKCDQFAVEYEPVLIETLVWMDPSFVCLKIGACP 486
QY 478 SAHKPLGTEKCIWGPSTYWCNTETAACNAVEHCKRHVMN 518
DB 487 SAHKPLGTEKCIWGPSTYWCNTETAACNAVEHCKRHVMN 527

RESULT 6
US-10-060-036-73
/ Sequence 73, Application US/10060036
/ Publication No. US20030073144A1
/ GENERAL INFORMATION:
/ APPLICANT: Benson, Darin R.
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Lodes, Michael J.
/ APPLICANT: Persing, David H.
/ APPLICANT: Hepler, William T.
/ APPLICANT: Jjiang, Yugu
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ FILE REFERENCE: 210121.566
/ CURRENT APPLICATION NUMBER: US/10/060,036
/ CURRENT FILING DATE: 2002-01-30
/ NUMBER OF SEQ ID NOS: 4560
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 73
/ LENGTH: 527
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-060-036-73

Query Match          99.6%; Score 2747.5; DB 14; Length 527;
Best Local Similarity 99.4%; Pred. No. 5.4e-236;
```

Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```

QY 1 LASLIGALAGPVGLKECTRGSAVWCQNYKTASDCGAVGHCIQTWNKPTVSLPCDIC 60
DB 7 LASLIGALAGPVGLKECTRGSAVWCQNYKTASDCGAVGHCIQTWNKPTVSLPCDIC 66
QY 61 KDVTAAAGDMLKDNATEEELIVLEKTCMDLPRKNSASCKEIVDSYLPVILDIIGEMS 120
DB 67 KDVTAAAGDMLKDNATEEELIVLEKTCMDLPRKNSASCKEIVDSYLPVILDIIGEMS 126
QY 121 RRGVCSALNLCESLQGHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLLYPDQGR 180
DB 127 RRGVCSALNLCESLQGHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLLYPDQGR 186
QY 181 SKQPKONGDVCCDIOMVTDIOTAVRTNSTFVQALVEHKECDRLGPMADICKNYIS 240
DB 187 SKQPKONGDVCCDIOMVTDIOTAVRTNSTFVQALVEHKECDRLGPMADICKNYIS 246
QY 241 QYSEIAIQMMHMM--OPKEICALVGFCEVKEMPMQTLVPAKASKNYIPALEVEPIK 297
DB 247 QYSEIAIQMMHMMQDQPKKEICALVGFCEVKEMPMQTLVPAKASKNYIPALEVEPIK 306
QY 298 KHEVPAKSDVYCEVCEFLVEVTKLIDNNKTEKEIIDAFDKMSCKLPSLSECEVVDYTS 357
DB 307 KHEVPAKSDVYCEVCEFLVEVTKLIDNNKTEKEIIDAFDKMSCKLPSLSECEVVDYTS 366
QY 358 YGSSILSILLEEVSPELVCSMLHLCSGTRLPALTVHTOPKGGFCVCKLVGLDRLNL 417
DB 367 YGSSILSILLEEVSPELVCSMLHLCSGTRLPALTVHTOPKGGFCVCKLVGLDRLNL 426
QY 418 EKSTKOEIIAALKEKCSFLPDPYQKCDQFVAEYEPVILIEIVEMDPSFVCLIKIGACP 477
DB 427 EKSTKOEIIAALKEKCSFLPDPYQKCDQFVAEYEPVILIEIVEMDPSFVCLIKIGACP 486
QY 478 SAKHPLGTGKCIWGPSYWCQNTETAOCNAVEHCKRHVNN 518
DB 487 SAKHPLGTGKCIWGPSYWCQNTETAOCNAVEHCKRHVNN 527

RESULT 7
US-09-767-007A-2
; Sequence 2, Application US/09767007A
; Patent No. US20020077275A1
; GENERAL INFORMATION:
; APPLICANT: John S. O'Brien
; APPLICANT: Yasuo Kishimoto
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
; FILE REFERENCE: MYELOS-2DC1C1
; CURRENT APPLICATION NUMBER: US/09/767,007A
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 08/958,970
; PRIOR FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 523
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-767-007A-2

```

Query Match 98.9%; Score 2727.5; DB 9; Length 523;
Best Local Similarity 99.2%; Pred. No. 3.3e-224;
Matches 514; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 LASLIGALAGPVGLKECTRGSAVWCQNYKTASDCGAVGHCIQTWNKPTVSLPCDIC 60
DB 7 LASLIGALAGPVGLKECTRGSAVWCQNYKTASDCGAVGHCIQTWNKPTVSLPCDIC 66
QY 61 KDVTAAAGDMLKDNATEEELIVLEKTCMDLPRKNSASCKEIVDSYLPVILDIIGEMS 120
DB 67 KDVTAAAGDMLKDNATEEELIVLEKTCMDLPRKNSASCKEIVDSYLPVILDIIGEMS 126

```

```

QY 121 RRGVCSALNLCESLQGHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLLYPDQGR 180
DB 127 RRGVCSALNLCESLQGHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLLYPDQGR 186
QY 181 SKQPKONGDVCCDIOMVTDIOTAVRTNSTFVQALVEHKECDRLGPMADICKNYIS 240
DB 187 SKQPKONGDVCCDIOMVTDIOTAVRTNSTFVQALVEHKECDRLGPMADICKNYIS 246
QY 241 QYSEIAIQMMHMMQPKKEICALVGFCEVKEMPMQTLVPAKASKNYIPALEVEPIK 300
DB 246 QYSEIAIQMMHMMQPKKEICALVGFCEVKEMPMQTLVPAKASKNYIPALEVEPIK 305
QY 301 VPAKSDVYCEVCEFLVEVTKLIDNNKTEKEIIDAFDKMSCKLPSLSECEVVDYTS 360
DB 306 VPAKSDVYCEVCEFLVEVTKLIDNNKTEKEIIDAFDKMSCKLPSLSECEVVDYTS 365
QY 361 SLSILLEEVSPELVCSMLHLCSGTRLPALTVHTOPKGGFCVCKLVGLDRLNL 420
DB 366 SLSILLEEVSPELVCSMLHLCSGTRLPALTVHTOPKGGFCVCKLVGLDRLNL 425
QY 421 STKOEIIAALKEKCSFLPDPYQKCDQFVAEYEPVILIEIVEMDPSFVCLIKIGACPSAH 480
DB 426 STKOEIIAALKEKCSFLPDPYQKCDQFVAEYEPVILIEIVEMDPSFVCLIKIGACPSAH 485
QY 481 KPLGTGKCIWGPSYWCQNTETAOCNAVEHCKRHVNN 518
DB 486 KPLGTGKCIWGPSYWCQNTETAOCNAVEHCKRHVNN 523

```

```

RESULT 8
US-09-978-418-40
; Sequence 40, Application US/09978418
; Publication No. US20030118997A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephan
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 142 US5, REG
; CURRENT APPLICATION NUMBER: US/09/978,418
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/311,305
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/314,734
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/318,204
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/326,470
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: JPatent
; SEQ ID NO 40
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-418-40

```

Query Match 87.7%; Score 2419.5; DB 10; Length 479;
Best Local Similarity 99.4%; Pred. No. 8.3e-207;
Matches 463; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```

QY 1 LASLIGALAGPVGLKECTRGSAVWCQNYKTASDCGAVGHCIQTWNKPTVSLPCDIC 60
DB 7 LASLIGALAGPVGLKECTRGSAVWCQNYKTASDCGAVGHCIQTWNKPTVSLPCDIC 66
QY 61 KDVTAAAGDMLKDNATEEELIVLEKTCMDLPRKNSASCKEIVDSYLPVILDIIGEMS 120
DB 67 KDVTAAAGDMLKDNATEEELIVLEKTCMDLPRKNSASCKEIVDSYLPVILDIIGEMS 126
QY 121 RRGVCSALNLCESLQGHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLLYPDQGR 180
DB 127 RRGVCSALNLCESLQGHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLLYPDQGR 186
QY 181 SKQPKONGDVCCDIOMVTDIOTAVRTNSTFVQALVEHKECDRLGPMADICKNYIS 240

```

Db 187 SKPGKNDGVCQDCIQWTDIQAVRTNSFVQALVHEVKEECRCRLPGMADICKNYIS 246
Qy 241 QYSEIAIOMMMHMM---QPKETCALVGFCDVEKEMQOTLVPAKVASKNVITPALEIVEIK 297
Db 247 QYSEIAIOMMMHMMQDOQPEKICALVGFCDVEKEMQOTLVPAKVASKNVITPALEIVEIK 306
Qy 298 KHEVPAKSDVYCEVCEPLVKEVTKLIDNNKTEKEILDAFDKMSKFLPKSLSECOEVDVT 357
Db 307 KHEVPAKSDVYCEVCEPLVKEVTKLIDNNKTEKEILDAFDKMSKFLPKSLSECOEVDVT 366
Qy 358 YGSSILSLLEEVSEPLVCSMLHLCSGTRLPALTVHTVQPKDGFCEVCKLVGYLDRL 417
Db 367 YGSSILSLLEEVSEPLVCSMLHLCSGTRLPALTVHTVQPKDGFCEVCKLVGYLDRL 426
Qy 418 EKNSTKOEIILAEKGSFLPPYQKQCDQFAVEEPVILIELVAV 463
Db 427 EKNSTKOEIILAEKGSFLPPYQKQCDQFAVEEPVILIELVAV 472

RESULT 9

US-10-205-194-176
; Sequence 176, Application US/10205194
; Publication No. US20030134301A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pimock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018201
; CURRENT APPLICATION NUMBER: US/10/205,194
; PRIOR FILING DATE: 5200-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 176
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: Prosaposin
US-10-205-194-176

Query Match 70.6%; Score 1947; DB 14; Length 554;

Best Local Similarity 65.6%; Pred. No. 1,4e-164;
Matches 360; Conservative 77; Mismatches 80; Indels 32; Gaps 3;

Qy 1 LASLIGALAGPVGLKCTRGSAVWCQNVKTASQCAVKGCLQTVVNNKPTVYKSLPDCIC 60
Db 7 LASLIVTLVLTSPVQDPKICSGGSAVVCBVDKTAVCRVAKHQQVWWSKPTAKSLPCIC 66
Qy 61 KDVTAAADMLKDNATEEELIYVLEKTCMDLPEKPNMSASCKEIVDSYLPVILDIKSGMS 120
Db 67 KTVVTEAGNLKDNATEEELIYVLEKTCMDLPEKPNMSASCKEIVDSYLPVILDIKSGMS 126
Qy 121 RPEVCSALNTCESLQKLAELNHOKLESNKIPILDMTEVAVPMANIPLLLYPOQDPR 180
Db 127 NPEVCSALNTCESLQKLAELNHOKLESNKIPILDMTEVAVPMANIPLLLYPOQDPR 185
Qy 181 SKPQPKNDGVCQDCIQWTDIQAVRTNSFVQALVHEVKEECRCRLPGMADICKNYIS 240
Db 186 SQPQPKNEEDVCQDCIKVTDIQAVRTNSFVQALVHEVKEECRCRLPGMADICKNYIS 245
Qy 241 QYSEIAIOMMMHMMQPEKICALVGFCDVEKEMQOTLVPAKVASKNVITPALEIVEIK 300
Db 246 QYSEIAIOMMMHMMQPEKICALVGFCDVEKEMQOTLVPAKVASKNVITPALEIVEIK 305
Qy 301 VPAKSDVYCEVCEPLVKEVTKLIDNNKTEKEILDAFDKMSKFLPKSLSECOEVDVYGS 360
Db 306 IQAQNVIFCQVQCLVWRKLSLIIINNATEELIKGLSKACSLIPAPASTKQCEVLVTRGP 365

Qy 361 SILSLLEEVSEPLVCSMLHLCSG-----TRLP-----A 389
Db 366 SLDLVLEHENVNPLCGVISLCSANPVLGTLBPAAAIYSAUPKPEAPPKQPEPKOSA 425
Qy 390 LTVHTVQPKDGFCEVCKLVGYLDRLNLEKSTKOEIILAEKGSFLPPYQKQCDQFV 449
Db 426 LRAHVPPQKNGGFCEVCKLVGYLDRLNLEKSTKOEIILAEKGSFLPPYQKQCDQFV 485
Qy 450 AEVEPVLIELVEVMDSPFVCLKTGACPSAKPLGTBEKCIWGSYWCQNTETAQCNV 509
Db 486 AEVEPVLIELVEVMDSPFVCLKTGACPSAKPLGTBEKCIWGSYWCQNTETAQCNV 545
Qy 510 EHCRRHYWN 518
Db 546 DHCRRHYWN 554

RESULT 10

US-10-276-162-1
; Sequence 1, Application US/10276162
; Publication No. US20030215822A1
; GENERAL INFORMATION:
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: YAO, Monique G.
; APPLICANT: BRUNS, Christopher M.
; APPLICANT: YUE, Henry
; APPLICANT: DELEGANE, Angelo M.
; APPLICANT: HAFALIA, April
; APPLICANT: PATTERSON, Chandra
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Daniel B.
; APPLICANT: LAU, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: JACKSON, Jennifer L.
; APPLICANT: LU, Dying Aina M.
; APPLICANT: BATRA, Sajeev
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: REDDY, Roopa
; APPLICANT: AZIMZAI, Yalda
; TITLE OF INVENTION: SECRETED PROTEINS
; FILE REFERENCE: PI-0071 USN
; CURRENT APPLICATION NUMBER: US/10/276,162
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US01/11861
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/197,854
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/202,373
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/205,899
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 60/210,155
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/209,401
; PRIOR FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030215822A1 7473577CD1
US-10-276-162-1

Query Match 43.1%; Score 1189; DB 15; Length 521;
Best Local Similarity 44.1%; Pred. No. 4,9e-97;
Matches 234; Conservative 98; Mismatches 157; Indels 42; Gaps 10;

```

QY 1 LASLIGALAGPVLGLKECTRGSAVMCONVKTASDGAVKHCLQTWANKPTVKSICDTC 60
DB 8 LPSLIGATRASPSGPECAKSTVWCODIQTAARCAVGTCCGAANNKPTASLPCDVC 67
QY 61 KDVTYTAGDMUKDNATEEELIVYEKTCDWLPRKNNASCKEIVDSYLPVILDIIGEM- 119
DB 68 QDIAAAGNGINPATESDITALVMKTCEWLPSQESAGCKMMVDAHSSAILSLMUGARD 127
QY 120 SRPEVCSALNLCESLQKHLAEINHOKQLESNKIPELDMTEVAVPMMANIPILLYPODGP 179
DB 128 SAPAQVCTALSLCEPILGRHLATL-----RPLSKEDTEBAVAFPMAGPLTFHPRQAP 179
QY 180 RSKPQPKDNGVDCDICIOMVTDIQTAVRTNSTFVQALVEHVEKCECDRLPGMADICKNYI 239
DB 180 -----EGALCODCVQRVSRLOEAVRSNLTADL---NIQRCESLPGIHLAVLCNLY 228
QY 240 SQYSEIALQMMHMQPEIKCALVGFCDVEKEMQTLVPAK---VASKNVIPLBELVEPI 296
DB 229 FQFVFPADQALRLPPELCEKRGKGFCEELG-----APARLTQVYAMDGVPSELGLPR 281
QY 297 KKHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMCKSLPKSLSECCQEVND 356
DB 282 KQSEMOMKAGVTCVCEMANNVQKLDHWMNSSELMITHALERCSVMPASITKECIIIVD 341
QY 357 TYGSSILSLLEEVSEPVCSMLHLCSGTRLPALTYH-----VTOPK-----DGFCEVC 406
DB 342 TYSPSLVQ-LVAKITPEKCKFTRLC-GNRRARAVHDAVAVPSPEDMAENQSFQNGC 399
QY 407 KKLVGYLDRNLEKNSTKOEILALKEGCSFLPDYQKODQFVAEYEPVILIVEVMP 466
DB 400 KRLITVSSHLSKSTKRDILVAFKGGCSILPLPYMIQCKHFTVQTEPVLISLKMMDP 459
QY 467 SFVCLKIGACPSAHKRLGTETKCIWGPSTWCONTEETAACQNAVEHCKRHAW 517
DB 460 VAVCKKVGACHGPRFTPLGTQDQALGSPFWRCSQGAALKCANVQHCQGHV 510

```

RESULT 11

```

US-09-833-245-903
; Sequence 903, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS46PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 903
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-903

```

```

Query Match 40.9%; Score 1127.5; DB 11; Length 531;
Best Local Similarity 42.9%; Pred. No. 1.5e-91;
Matches 228; Conservative 99; Mismatches 161; Indels 43; Gaps 11;
QY 1 LASLIGALAGPVLGLKECTRGSAVMCONVKTASDGAVKHCLQTWANKPTVKSICDTC 60
DB 8 LPSLIGATRASPSGPECAKSTVWCODIQTAARCAVGTCCGAANNKPTASLPCDVC 67
QY 61 KDVTYTAGDMUKDNATEEELIVYEKTCDWLPRKNNASCKEIVDSYLPVILDIIGEM- 119
DB 68 QDIAAAGNGINPATESDITALVMKTCEWLPSQESAGCKMMVDAHSSAILSLMUGARD 127

```

```

QY 120 SRPEVCSALNLCESLQKHLAEINHOKQLESNKIPELDMTEVAVPMMANIPILLYPODGP 179
DB 128 SAPAQVCTALSLCEPILGRHLATL-----RPLSKEDTEBAVAFPMAGPLTFHPRQAP 179
QY 180 RSKPQPKDNGVDCDICIOMVTDIQTAVRTNSTFVQALVEHVEKCECDRLPGMADICKNYI 239
DB 180 -----EGALCODCVQRVSRLOEAVRSNLTADL---NIQRCESLPGIHLAVLCNLY 228
QY 240 SQYSEIALQMMHMQPEIKCALVGFCDVEKEMQTLVPAK---VASKNVIPLBELVEPI 296
DB 229 FQFVFPADQALRLPPELCEKRGKGFCEELG-----APARLTQVYAMDGVPSELGLPR 281
QY 297 KKHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMCKSLPKSLSECCQEVND 356
DB 282 KQSEMOMKAGVTCVCEMANNVQKLDHWMNSSELMITHALERCSVMPASITKECIIIVD 341
QY 357 TYGSSILSLLEEVSEPVCSMLHLCSGTRLPALTYH-----VTOPK-----DGFCEVC 406
DB 342 TYSPSLVQ-LVAKITPEKCKFTRLC-GNRRARAVHDAVAVPSPEDMAENQSFQNGC 399
QY 407 KKLVGYLDRNLEKNSTKOEILALKEGCSFLPDYQKODQFVAEYEPVILIVEVMP 466
DB 400 KRLITVSSHLSKSTKRDILVAFKGGCSILPLPYMIQCKHFTVQTEPVLISLKMMDP 459
QY 467 SFVCLKIGACPSAHKRLGTETKCIWGPSTWCONTEETAACQNAVEHCKRHAW 517
DB 460 VAVCKKVGACHGPRFTPLGTQDQALGSPFWRCSQGAALKCANVQHCQGHV 509

```

RESULT 12

```

US-10-043-487-340
; Sequence 340, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypeptide
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 340
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Shigella flexneri
US-10-043-487-340

```

```

Query Match 39.6%; Score 1093; DB 14; Length 209;
Best Local Similarity 100.0%; Pred. No. 4.ee-89;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 145 OKQLESNKIPELDMTEVAVPMMANIPILLYPODGPGRKPKDNGVDCDICIOMVTDIQT 204
DB 1 OKQLESNKIPELDMTEVAVPMMANIPILLYPODGPGRKPKDNGVDCDICIOMVTDIQT 60
QY 205 AVRTNSTFVQALVEHVEKCECDRLPGMADICKNYISQYSEIALQMMHMQPEIKCALVGF 264
DB 61 AVRTNSTFVQALVEHVEKCECDRLPGMADICKNYISQYSEIALQMMHMQPEIKCALVGF 120
QY 265 CDEVKEMQTLVPAKASKNVILALBELVEPIKHEVPAKSDVYCEVCEFLVKEVTKLID 324
DB 121 CDEVKEMQTLVPAKASKNVILALBELVEPIKHEVPAKSDVYCEVCEFLVKEVTKLID 180
QY 325 NNKTEKEILDAFDMCKSLPKSLSECCOE 353
DB 181 NNKTEKEILDAFDMCKSLPKSLSECCOE 209

```

RESULT 13

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:26:23 ; Search time 20.661 Seconds

(without alignments)
2411.659 Million cell updates/secTitle: US-09-743-684a-1_COPY_7_524
Perfect score: 2759
Sequence: 1 LASLIGALAGPVGLKECT.....NTETAAQCNAYEHCRRHWN 518Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2747.5	99.6	527	1	SAHUP
2	1974	71.5	554	1	A28716
3	1915.5	69.4	557	2	JH0604
4	511	18.5	965	2	T00207
5	389	14.1	79	2	A49475
6	372	13.5	80	2	S02766
7	343	12.4	86	2	S21770
8	339.5	12.3	381	1	LNHUB
9	305	11.1	363	2	A29072
10	293	10.6	369	2	I46531
11	261	9.5	81	2	A32026
12	250	9.1	370	1	LNRBB
13	213	7.7	213	2	T46069
14	203.5	7.4	417	2	T48201
15	176.5	6.4	402	2	T15677
16	173.5	6.3	513	2	T09739
17	171.5	6.2	506	2	T09739
18	165.5	6.0	506	2	T07915
19	161.5	5.9	513	2	T11686
20	161	5.8	428	2	S47096
21	161	5.8	474	2	T12049
22	158.5	5.7	508	2	S19697
23	156	5.7	314	2	T15674
24	155.5	5.6	292	2	T14445
25	155.5	5.6	322	2	S41400
26	154	5.6	496	2	JS0732
27	153.5	5.6	409	2	JS0732
28	145.5	5.3	205	2	B89567
29	143.5	5.2	1175	2	S52417

oryzasin (EC 3.4.2
cyprosin (EC 3.4.2
gene 11-1 protein
probable aspartic
cag pathogenicity
hypothetical prote
aspartic proteins
cysteine-rich fibr
hypothetical prote
cag island protein
aspartic proteins
hypothetical prote
probable CDP-rich
hypothetical prote
resc8 - rat (fragm
protein FIN21.4 [1

ALIGNMENTS

RESULT 1

SAHUP
saposin precursor [validated] - human
N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component
ein (SAP); sphingolipid activator protein A2; sulfatide sulfatase activator protein
N:Contents: prosaposin; saposin A; saposin B; saposin C; saposin D
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1992 #sequence revision 17-Nov-1995 #text change 08-Dec-2000
C:Accession: JX0061; A57368; A42003; B42003; C42003; A30367; S34740; S36140; S36
0226; I37265; I37264
R:Nakano, T.; Sandhoff, K.; Stuenkel, J.; Christomanou, H.; Suzuki, K.
J. Biochem. 105, 152-154, 1989
A:Title: Structure of full-length cDNA coding for sulfatide activator, a Co-beta-glucosi
A:Reference number: JX0061; MUID:89255151; PMID:2498298
A:Accession: JX0061
A:Molecule type: mRNA
A:Residues: 1-527 <NMK>
A:Cross-references: GB:D00422; NID:g220063; PIDN:BAA0021.1; PID:g220064
A>Note: alternative splice form 1
A:Accession: A57368
A:Molecule type: mRNA
A:Residues: 1-259, 263-527 <NA2>
A:Cross-references: GB:J03015; GB:J03086; NID:g337755; PIDN:AAB5494.1; PID:g337756
A>Note: alternative splice form 2
R:Rotman, E.G.; Scheinker, V.; Grabowski, G.A.
Genomics 13, 312-318, 1992
A:Title: Structure and evolution of the human prosaposin chromosomal gene.
A:Reference number: A42003; MUID:92307663; PMID:1612590
A:Accession: A42003
A:Molecule type: DNA
A:Residues: 50-140 <ROR>
A:Cross-references: GB:M6181
A>Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIIP:107236)
A:Accession: B42003
A:Molecule type: DNA
A:Residues: 185-259, 263-276 <RO2>
A>Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIIP:107237)
A:Accession: C42003
A:Molecule type: DNA
A:Residues: 305-393 <RO3>
A>Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIIP:107238); sequence inc
A:Accession: D42003
A:Molecule type: DNA
A:Residues: 399-487 <RO4>
A>Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIIP:107239); sequence inc
R:Rotman, E.G.; Grabowski, G.A.
Genomics 5, 486-492, 1989
A:Title: Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphin
A:Reference number: A30367; MUID:90129043; PMID:2515150
A:Accession: A30367
A:Molecule type: mRNA
A:Residues: 1-259, 263-527 <RO5>

A/Cross-references: GB:J03077; NID:9183230; PIDN:AAA52560.1; PID:9183231
 A/Note: alternative splice form 2
 R/Hirata, M.; O'Brien, J.S.; Kishimoto, Y.; Galdzicka, M.; Fluharty, A.L.; Gims, E.L.; Arch. Biochem. Biophys. 304, 110-116, 1993
 A/Title: Isolation, characterization, and proteolysis of human prosaposin, the precursor A/Accession: S34740
 A/Molecule type: protein
 A/Residues: 17-24;165-172;180-189;301-305 <HR>
 R/Tymela, J.; Palmer, D.N.; Baumann, M.; Haltia, M. FEBS Lett. 330, 8-12, 1993
 A/Title: Storage of saposin A and D in infantile neuronal ceroid-lipofuscinosis. A/Reference number: S36140; MUID:93380576; PMID:8370464
 A/Accession: S36140
 A/Molecule type: protein
 A/Residues: 'XX', 62, 'X', 64-65, 'X', 67-79, 'X', 81-84 <TV>
 A/Note: saposin A
 A/Accession: S36141
 A/Molecule type: protein
 A/Residues: 'XXX', 413-414, 'X', 416-428, 'X', 430-434 <TV>
 A/Note: saposin D
 R/Holtschmidt, H.; Sandhoff, K.; Kwon, H.Y.; Harzer, K.; Nakano, T.; Suzuki, K. J. Biol. Chem. 266, 7556-7560, 1991
 A/Title: Sulfatide activator protein. Alternative splicing that generates three mRNAs at A/Reference number: S36988; MUID:91210267; PMID:2019586
 A/Accession: S36988
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-240, 'S', 242-259, 261-527 <HO>
 A/Cross-references: EMBL:M60255; NID:9337759; PIDN:AAA5594.1; PID:9337760
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
 A/Title: cerebroside sulfate activator protein mutant MU-9; corresponds to alternative sp A/Accession: S35989
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-240, 'S', 242-259, 261-527 <HO>
 A/Cross-references: EMBL:M60257; NID:9337764; PIDN:AAA5595.1; PID:9337765
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
 A/Title: cerebroside sulfate activator protein mutant MU-0; corresponds to alternative sp A/Accession: S36990
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-240, 'S', 242-259, 261-527 <HO>
 A/Cross-references: EMBL:M60258; NID:9337766; PIDN:AAA5596.1; PID:9337767
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991
 A/Title: cerebroside sulfate activator protein mutant MU-6; corresponds to alternative sp R/Kondoh, K.; Hiheno, T.; Sano, A.; Kakimoto, Y. Biochem. Biophys. Res. Commun. 181, 286-292, 1991
 A/Title: Isolation and characterization of prosaposin from human milk. A/Reference number: P80330; MUID:92068206; PMID:1958198
 A/Accession: P80330
 A/Molecule type: protein
 A/Residues: 17-24, 'X', 26 <KON>
 A/Experimental source: milk
 R/Kretz, K.A.; Carson, G.S.; Morimoto, S.; Kishimoto, Y.; Fluharty, A.L.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 87, 2541-2544, 1990
 A/Title: Characterization of a mutation in a family with saposin B deficiency: a glycosy A/Reference number: A35985; MUID:90207231; PMID:2320574
 A/Accession: A35985
 A/Molecule type: mRNA
 A/Residues: 213-221 <KRE>
 A/Cross-references: GB:M32221
 A/Accession: B35985
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-259, 263-527 <KR>
 A/Cross-references: GB:M32221; NID:9337761; PIDN:AAA60303.1; PID:9337762
 A/Experimental source: lymphoblast
 A/Accession: C35985
 A/Molecule type: mRNA
 A/Residues: 213-216, 'I', 218-221 <KR>
 A/Note: sequence from patients with activator-deficient metachromatic leukodystrophy; ch R/Fuerst, W.; Schubert, J.; Machleidt, W.; Meyer, H.E.; Sandhoff, K.

Eur. J. Biochem. 192, 709-714, 1990
 A/Title: The complete amino-acid sequences of human ganglioside GM2 activator protein an A/Reference number: S13195; MUID:9106165; PMID:2209618
 A/Accession: S13196
 A/Molecule type: protein
 A/Residues: 195-259, 263-277 <FUE>
 R/Morimoto, S.; Martin, B.M.; Yamamoto, Y.; Kretz, K.A.; O'Brien, J.S.; Kishimoto, Y. Proc. Natl. Acad. Sci. U.S.A. 86, 3389-3393, 1989
 A/Title: Saposin A: second cerebrosidase activator protein. A/Reference number: A32784; MUID:89240739; PMID:2717620
 A/Accession: A32784
 A/Molecule type: protein
 A/Residues: 60-84;86-107;109-119;125-134 <MOR>
 R/O'Brien, J.S.; Kretz, K.A.; Dewji, N.; Wenger, D.A.; Each, F.; Fluharty, A.L. Science 241, 1098-1101, 1988
 A/Title: Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic A/Reference number: A41240; MUID:88321660; PMID:2842863
 A/Accession: A41240
 A/Molecule type: mRNA
 A/Residues: 'GSSR', 18-259, 263-299, 'D', 301-302, 'D', 304-527 <OAB>
 A/Cross-references: GB:J03086
 R/Dewji, N.N.; Wenger, D.A.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 84, 8652-8656, 1987
 A/Title: Nucleotide sequence of cloned cDNA for human sphingolipid activator protein 1 p A/Reference number: S02289; MUID:88068647; PMID:2825202
 A/Accession: S02289
 A/Status: significant sequence differences
 A/Molecule type: mRNA
 A/Cross-references: EMBL:J03015
 A/Note: this sequence corrected by A41240
 A/Note: part of this sequence, including the amino end of the mature protein, was determ R/Kleinhardt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 369, 1361-1365, 1988
 A/Title: Complete amino-acid sequence of the naturally occurring A(2) activator protein 1 A/Reference number: S02028; MUID:89207118; PMID:3242555
 A/Accession: S02028
 A/Molecule type: protein
 A/Residues: 195-259, 263-276 <KLE>
 R/Fuerst, W.; Machleidt, W.; Sandhoff, K. Biol. Chem. Hoppe-Seyler 369, 317-328, 1988
 A/Title: The precursor of sulfatide activator protein is processed to three different pr A/Reference number: S00813; MUID:89000190; PMID:3048308
 A/Accession: S00813
 A/Molecule type: protein
 A/Residues: 410-487 <PU>
 R/Kleinhardt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 368, 1571-1578, 1987
 A/Title: Complete amino-acid sequence and carbohydrate content of the naturally occurin A/Reference number: S00226; MUID:88163077; PMID:3442600
 A/Accession: S00226
 A/Molecule type: protein
 A/Residues: 314-393 <KL>
 R/Vaccaro, A.M.; Salvio, R.; Barca, A.; Tatti, M.; Ciaffoni, F.; Maras, B.; Siciliano, U. Biol. Chem. 270, 9953-9960, 1995
 A/Title: Structural analysis of saposin C and B. Complete localization of disulfide bridg A/Reference number: A57297; MUID:95247790; PMID:7730378
 A/Contents: annotation; disulfide bonds; glycosylation
 R/Holtschmidt, H.; Sandhoff, K.; Fuerst, W.; Kwon, H.Y.; Schnabel, D.; Suzuki, K. FEBS Lett. 280, 267-270, 1991
 A/Title: The organization of the gene for the human cerebrosidase sulfate activator protei A/Reference number: 137264; MUID:91192146; PMID:2013321
 A/Accession: 137265
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 59-125 <RES>
 A/Cross-references: EMBL:X57107; NID:930234; PIDN:CAA40391.1; PID:930235
 A/Accession: 137264
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 307-516 <RE>
 A/Cross-references: EMBL:X57108; NID:930232; PIDN:CAA40392.1; PID:91565257
 A/Note: sequence revised relative to PID:930233 (corrected coding region)
 C/Genetics:

A:Gene: GDB:PSAP; GLBA
A:Cross-references: GDB:120366; OMIM:176801
A:Map position: 10q22.1-10q22.1
A:Introns: 83/3; 338/3; 401/1; 453/3; 480/3
A>Note: defects in this gene may cause variant Gaucher disease, variant Tay-Sachs disease
A>Note: list of introns is incomplete

Query Match 99.6%; Score 2747.5; DB 1; Length 527;
Best Local Similarity 99.4%; Pred. No. 176-173;
Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```
QY 1 LASLGAALAGPVGLKECTRGSAVWCONVKTASDCGAKHCLQTVNMKPTVKSJPCIC 60
DB 7 LASLGAALAGPVGLKECTRGSAVWCONVKTASDCGAKHCLQTVNMKPTVKSJPCIC 66
QY 61 KDVTAAAGDMLKDNATEBEILVLEKTCMDLPKPMMSASCKEIVDSYLPVLIIDIKGMS 120
DB 67 KDVTAAAGDMLKDNATEBEILVLEKTCMDLPKPMMSASCKEIVDSYLPVLIIDIKGMS 126
QY 121 RPEVCSALNLCESLOKHLAELNHOKLESNKIPELDMTEVAVAPPMANIPILLYPQDGR 180
DB 127 RPEVCSALNLCESLOKHLAELNHOKLESNKIPELDMTEVAVAPPMANIPILLYPQDGR 186
QY 181 SKPOKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVHVKEECRLGFGMADICKNYIS 240
DB 187 SKPOKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVHVKEECRLGFGMADICKNYIS 246
QY 241 QYSEIAIOMMMHMOOQPEKICALVGFCDVEKEMPMQTLVPAKVASKNVIPLALVEPIK 297
DB 247 QYSEIAIOMMMHMOOQPEKICALVGFCDVEKEMPMQTLVPAKVASKNVIPLALVEPIK 306
QY 298 KHEVPAKSDVYCEVEFLVKEVTKLIDNNKTEKELIDAFDMKCSLPSISECOEVDVT 357
DB 307 KHEVPAKSDVYCEVEFLVKEVTKLIDNNKTEKELIDAFDMKCSLPSISECOEVDVT 366
QY 358 YGSSILSTILLEVEBPVLCVSMHLCSGTRLPALTVHTVQPKDGFCEVCKLVGILDRNL 417
DB 367 YGSSILSTILLEVEBPVLCVSMHLCSGTRLPALTVHTVQPKDGFCEVCKLVGILDRNL 426
QY 418 EKNSIKOELIALAEKGSFLLPDPYOKQCDQFAVEYEPVILIEIVWMDPSFVCLKIGACP 477
DB 427 EKNSIKOELIALAEKGSFLLPDPYOKQCDQFAVEYEPVILIEIVWMDPSFVCLKIGACP 486
QY 478 SAHKPLGTETKICWGPSYWCNTETAAOCNAVEHKKRHVN 518
DB 487 SAHKPLGTETKICWGPSYWCNTETAAOCNAVEHKKRHVN 527
```

RESULT 2

A28716

saposin precursor - rat

N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfate
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A28716

R:Collard, M.W.; Sylvestre, S.R.; Tsunrta, J.K.; Griswold, M.D.
Biochemistry 27, 4557-4564, 1988
A:Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat B

A:Reference number: A28716; MUID:89000647; PMID:3048385
A:Accession: A28716

A:Molecule type: mRNA
A:Residues: 1-554 <C>
A:Cross-references: GB:M19936; NID:9206904; PID:AAA42136.1; PID:9206905

A>Note: parts of this sequence, including the amino end of the mature protein, were de
C:Function:

A:Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc
A:Pathway: sphingolipid catabolism
A>Note: saposin A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy
A>Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsul
A>Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiester
C:Superfamily: saposin; saposin repeat homology

C:Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph

F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-554/Product: prosaposin #status predicted <PRO>
F:55-148/Domain: saposin repeat homology <SAP1>
F:160-143/Product: saposin A #status predicted <SAPA>
F:189-280/Domain: saposin repeat homology <SAP2>
F:194-273/Product: saposin B #status predicted <SAB1>
F:306-397/Domain: saposin repeat homology <SAP3>
F:310-522/Product: saposin C #status predicted <SAPC>
F:431-522/Domain: saposin repeat homology <SAP4>
F:437-514/Product: saposin D #status predicted <SAPD>
F:63-138,66-132,94-106,439-512,442-506,470-481/Dissulfide bonds: #status predicted
F:80,214,331,456/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:197-270,200-264,229-240,314-387,317-381,345-356/Dissulfide bonds: #status predicted

Query Match 71.5%; Score 1974; DB 1; Length 554;
Best Local Similarity 66.1%; Pred. No. 166-122;
Matches 363; Conservative 77; Mismatches 77; Indels 32; Gaps 3;

```
QY 1 LASLGAALAGPVGLKECTRGSAVWCONVKTASDCGAKHCLQTVNMKPTVKSJPCIC 60
DB 7 LASLGAALAGPVGLKECTRGSAVWCONVKTASDCGAKHCLQTVNMKPTVKSJPCIC 66
QY 61 KDVTAAAGDMLKDNATEBEILVLEKTCMDLPKPMMSASCKEIVDSYLPVLIIDIKGMS 120
DB 67 KDVTAAAGDMLKDNATEBEILVLEKTCMDLPKPMMSASCKEIVDSYLPVLIIDIKGMS 126
QY 121 RPEVCSALNLCESLOKHLAELNHOKLESNKIPELDMTEVAVAPPMANIPILLYPQDGR 180
DB 127 RPEVCSALNLCESLOKHLAELNHOKLESNKIPELDMTEVAVAPPMANIPILLYPQDGR 185
QY 181 SKPOKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVHVKEECRLGFGMADICKNYIS 240
DB 186 SKPOKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVHVKEECRLGFGMADICKNYIS 245
QY 241 QYSEIAIOMMMHMOOQPEKICALVGFCDVEKEMPMQTLVPAKVASKNVIPLALVEPIK 300
DB 246 QYSEIAIOMMMHMOOQPEKICALVGFCDVEKEMPMQTLVPAKVASKNVIPLALVEPIK 305
QY 301 VPASDVYCEVEFLVKEVTKLIDNNKTEKELIDAFDMKCSLPSISECOEVDVTYGS 360
DB 306 VPASDVYCEVEFLVKEVTKLIDNNKTEKELIDAFDMKCSLPSISECOEVDVTYGS 365
QY 361 SILSTILLEVEBPVLCVSMHLCSG-----TRLP-----A 389
DB 366 SILSTILLEVEBPVLCVSMHLCSG-----TRLP-----A 425
QY 390 LTVHTVQPKDGFCEVCKLVGILDRNLKNSKQELIALAEKGSFLLPDPYOKQCDQFV 449
DB 426 LRAHVPQKNGFCEVCKLVGILDRNLKNSKQELIALAEKGSFLLPDPYOKQCDQFV 485
QY 450 AVEYEPVILIEIVWMDPSFVCLKIGACPSAHKPLGTETKICWGPSYWCNTETAAOCNAV 509
DB 486 AVEYEPVILIEIVWMDPSFVCLKIGACPSAHKPLGTETKICWGPSYWCNTETAAOCNAV 545
QY 510 EHCKRHVN 518
DB 546 EHCKRHVN 554
```

RESULT 3

JH0604

saposin precursor - mouse

N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfate
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JH0604

R:Tsuada, M.; Sakiyama, T.; Endo, H.; Kitagawa, T.
Biochem. Biophys. Res. Commun. 184, 1266-1272, 1992
A:Title: The primary structure of mouse saposin.
A:Reference number: JH0604; MUID:92272718; PMID:1590788

A:Accession: JH0604
 A:Molecule type: mRNA
 A:Residues: 1-557 <TSU>
 A:Cross-references: GB:S36200; NID:G249386; PIDN:AA22175.1; PID:G249387
 A:Experimental source: liver
 C:Function:
 A:Description: saposin bind sphingolipids, form hydrophilic complexes and make them acc
 A:Pathway: sphingolipid catabolism
 A>Note: saposin A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy
 A>Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsulfa
 C:Superfamily: saposin; saposin repeat homology
 C:Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-557/Product: saposin repeat homology <SAP>
 F:55-143/Domain: saposin repeat homology <SAP1>
 F:56-143/Product: saposin repeat homology <SAP1>
 F:189-283/Domain: saposin repeat homology <SAP2>
 F:189-283/Product: saposin repeat homology <SAP2>
 F:309-400/Domain: saposin repeat homology <SAP3>
 F:309-400/Product: saposin repeat homology <SAP3>
 F:440-517/Domain: saposin repeat homology <SAP4>
 F:440-517/Product: saposin repeat homology <SAP4>
 F:63-138/Domain: saposin repeat homology <SAPD>
 F:63-138/Product: saposin repeat homology <SAPD>
 F:80-214,334,379,459/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 69.4%; Score 1915.5; DB 1; Length 557;
 Best Local Similarity 63.5%; Pred. No. 1.2e-118;
 Matches 350; Conservative 77; Mismatches 89; Indels 35; Gaps 3;

2 ASLIGALGPGVIGLKECTRGSAVWCQNVKTASDCGAVHCLQTVWNNKPTVSLPCDICK 61
 8 ASLIGALGPGVIGLKECTRGSAVWCQNVKTASDCGAVHCLQTVWNNKPTVSLPCDICK 67
 62 DVTTAGMMLKONATEEELIYVLEKTCMDLPRKNNASCKEIVDSYLVPIIDITKESNR 121
 68 TVTTAGMMLKONATEEELIYVLEKTCMDLPRKNNASCKEIVDSYLVPIIDITKESNR 127
 122 PEVGSALMLCSLQGHLELHNRQLESNKIPBLDMTEVAPFMANIPLLIYPODGRS 181
 128 PEVGSALMLCSLQGHLELHNRQLESNKIPBLDMTEVAPFMANIPLLIYPODGRS 186
 182 KPQKNDGVCODCIOMVTDIQTAVRINSTPQALVHVEKCECDLPGCADICKNYISQ 241
 187 QPQKNDGVCODCIOMVTDIQTAVRINSTPQALVHVEKCECDLPGCADICKNYISQ 246
 242 YSEIATQMMHMH--QPRKICLVGFCDEVEKEMQVTLVPAKVASKNVYIPALIEVEPIK 298
 247 YSEIATQMMHMH--QPRKICLVGFCDEVEKEMQVTLVPAKVASKNVYIPALIEVEPIK 306
 299 HEVPAKSPVCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSECEQVVDY 358
 307 NUVQAHNVITLCOTCOFVNNKSEELVNNATEELVGLSNACGLPDPARTKCEVVGTF 366
 359 GSSILSILLESVPVLCVSMHLCSG-----TRL 387
 367 GPSLIDIFIHVNNPSSLCGVIGLCAARELVLEALQAPAPVSLKLEPPPPQAPQKQ 426
 388 PALTVHTQPDGFCFCEVCKLVGLVLDNLEKSTKQELIILAELKSGFLPDPYQKQCD 447
 427 SALPAHVPQKNGFCEVCKLVGLVLDNLEKSTKQELIILAELKSGFLPDPYQKQCD 466
 448 FVAEYEVLLIETLVVNDPSPVLCIKIGACPSAHNPILGTEKCIWGSRYNQNTETAQCN 507
 487 FVAEYEVLLIETLVVNDPSPVLCIKIGACPSAHNPILGTEKCIWGSRYNQNTETAQCN 546
 QY 508 AVEHCKRHVNN 518
 Db 547 AVDHCKRHVNN 557

RESULT 4
 T00207

P109 protein - silkworm
 C:Species: Bombyx mori (silkworm)
 C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
 C:Accession: T00207
 R:Ramdun, J.; Chang, P.K.; Li, H.; Natori, M.
 Gene 212, 287-293, 1998
 A:Title: Molecular cloning of a cDNA encoding a silkworm protein which contains the cons
 A:Reference number: Z14124; MUID:96278844; PMID:9611271
 A:Accession: T00207
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-965 <TAN>
 A:Cross-references: EMBL:AB008449; NID:G2575864; PIDN:BA23126.1; PID:G2575865
 C:Superfamily: saposin repeat homology
 F:778-870/Domain: saposin repeat homology <SAP>

Query Match 18.5%; Score 511; DB 2; Length 965;
 Best Local Similarity 24.6%; Pred. No. 7.9e-26;
 Matches 139; Conservative 104; Mismatches 229; Indels 94; Gaps 18;

13 VLGLKECTRGSAVWCQNVKTASDCGAVHCLQTVWNNK--PTVKSIPCDICDVTAGD 69
 170 LKGSRTGWSPSTWCSNFTGRCCNAPHCINRWKMTFPEDDNDCQICIDMWQARD 229
 70 MLKONATEEELIYVLEKTCMDLPRKNNASCKEIVDSYLVPIIDITKESNRPGVCSAL 129
 230 QLOSNETQDIKEVFEKSGCKLPIKRVAGCKMLDEFVELIETLASNN--PQAVCSVA 288
 130 NLCS--LQKHLELHNRQLES-----NKIPBLDMTEVAPFMA----- 167
 289 GLCNNAKIDPLVDVNAQBELRAGCYNCCQTVGVARKKEDETKEDFLVGLLQVGRNMS 348
 168 --NIPLLYP-----QD-----PRSKPQKNDG 190
 349 LSDSCMLIFKYENILEAVKNDLPEGICHVSQGSYKFAHNDFTPEQWQYSATD 408
 191 V--QODICQWTDIQRAVRINSTPQALVHVEKCECDLPGCADICKNYISQSEIATQ 249
 409 VPCEFEQVLKRLDVLVANTTELE--FYKVLQGLCKQKQK--FKDECHLHAGQYVVIYNF 466
 250 MM--HQPKKEICLVGFCDEVEKEMQVTLV-----PAVLA-----SKNVIP 288
 467 LVSDLPARTCGMIGICGNLISAPISPLVARELVKVPKILGAEBSKIAVPLAKQMEP 526
 289 ALEIYE--PIKHEVPA--KSDYVCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPK 345
 527 ASAAVSVPLEEMFPAAPQSKAACAFQCFYHVLQVQSDTRTEDKVAVAQEAQDALPD 586
 346 SLSECEQVVDYGSILSILLESVPVLCVSMHLCSGTRLPALVHTQPDGFCFCEV 405
 587 ALNGECKEFVYQGSAAVIALVQELIDPASCPALQICQTE--EIRADVNSEKN--CPL 643
 406 CCKLVGYIDRLNENKSTKQELIILAELKSGFLPDPYQKQCDQFVAEYEVLLIETLVVND 465
 644 CLFAVEQLESVILKNRSEENIRKALDGLCTRLSQKLSCECIDFDYTSQGLVETLVADNN 703
 QY 466 PSFVCLTKIGAC--PSAHKPLIGTEKCI 490
 Db 704 AKETIVPLKCRDQLHDLKLTHTSSI 729

RESULT 5
 A49475
 Cerebroside sulfatase activator - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 17-May-1996
 C:Accession: A49475
 R:Stevens, R.L.; Fauli, K.F.; Conklin, K.A.; Green, B.N.; Fluharty, A.L.
 Biochemistry 32, 4051-4059, 1993
 A:Title: Porcine cerebroside sulfatase activator: further structural characterization and
 A:Reference number: A49475; MUID:93229506; PMID:8471613
 A:Status: Preliminary

A.Molecule type: protein
A.Residues: 1-79 <STB>
A.Experimental source: kidney
A.Note: sequence extracted from NCBI backbone (NCBIP:129597)
C:Superfamily: saposin; saposin repeat homology
F:1-79/Domain: saposin repeat homology <SAP>

Query Match 14.1% Score 389; DB 2; Length 79;
Best Local Similarity 88.6%; Pred. No. 3.8e-19;
Matches 70; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 189 GDVCCDCIQMTYDTIQAVTNSFVQALVEHYKECDRLGPGMADICKNYISQYSEIAIQ 248
DB 1 GDVCCDCIQMTYDTIQAVTNSFVQALVEHYKECDRLGPGMADICKNYISQYSEIAIQ 60
QY 249 MMMHMQPKDICKLVGFCDE 267
DB 61 MMMHMQPKDICKLVGFCDE 79

RESULT 6

S02766
pulmonary surfactant protein B precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 20-Aug-1999
C:Accession: S02766
R:Emile, P.A.; Shannon, J.M.; Mason, R.J.; Fisher, J.H.
Biochim. Biophys. Acta 994, 215-221, 1989
A>Title: cDNA and deduced amino acid sequence for the rat hydrophobic pulmonary surfactant
A:Reference number: S02766; MUID:89150284; PMID:2920185
A:Accession: S02766
A:Molecule type: mRNA
A:Residues: 1-376 <EMR>
A:Cross-references: EMBL:X14778; NID:957284; PIDN:CAA32885.1; PID:957285
C:Superfamily: pulmonary surfactant protein B; saposin repeat homology
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-190/Domain: propeptide #status predicted <PRO>
F:59-151/Domain: saposin repeat homology <SAP1>
F:190-277/Domain: saposin repeat homology <SAP2>
F:191-269/Product: pulmonary surfactant protein B #status predicted <MAT>
F:286-371/Domain: saposin repeat homology <SAP3>

Query Match 13.5% Score 372; DB 2; Length 376;
Best Local Similarity 22.7%; Pred. No. 3.5e-17;
Matches 115; Conservative 74; Mismatches 151; Indels 166; Gaps 17;

QY 5 LGALAGPYLGIKECTRGSAVWCQNVKTASDCAVAGHCLQTVNKPVTYSLPCDICKDV 64
DB 18 LGALAGPYLGIKECTRGSAVWCQNVKTASDCAVAGHCLQTVNKPVTYSLPCDICKDV 74
QY 65 TRAGMGLKDNATFEELVLEKTCWMLPKPMNSASCKEIVDSYLPVLIIDITIGMSRPG 124
DB 75 HLLTMTREDAQODTIRKFLQECEDILPKLKLVPKRCQVLDVLPVDFGQI-KPRA 133
QY 125 VCSALNLESLQKHLAELNHQQLBSNKLPELDMTEVAVPMANIPLLILPDGGRSKQ 184
DB 134 ICSHWGLC-----PL-----GQTKEQKE 153
QY 185 PKDNGDVCCDCIQMTYDTIQAVTNSFVQALVEHYKECDRLGPGMADICKNYISQYSE 244
DB 154 ----- 153
QY 245 IAIQMMHMQPKDICKLVGFCDEVKEMQTLVPAKVASKNVPALE---LVEP----- 295
DB 154 -----MLDAIPPLANKVLPALPGAFLLRPPHRTD 185
QY 296 IKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKIID-APDKMSKLPKSLSECEOV 354
DB 186 LSEQQLPIPLP-FCMLCRTLIKRVQAVI-----PKQVLAVANSQVCHVPLVVGICQCL 239
QY 355 VDTYSSILSLLEEVSPELVCSMLTCS-----GTRLPALTIVHTQ-PKDGFCFVCKK 408
DB 240 AERYIVLLDLALGRVPLQVCGVLVRCGTADAIGPALPALPLLEKMPLDQTECHFKCS 299

QY 409 LVGYLDRRLKRNSTKQIBLALBKGC-SFLPPYQKOCDFVAXEYEPVLIBVEWMDPS 467
DB 300 VI-----NQAMNTSQAMPQAMHQAQLRFWID--RQKEQVVEQMPOLLALVPSQDAH 352
QY 468 FVCELTGACPSAHKRLGTCKCIWGP 493
DB 353 TSCQALGVCEAPASPL-----QCFQTP 374

RESULT 7

S21770
saposin-C - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S21770
R:Sano, A.; Mizuno, T.; Kondoh, K.; Hineno, T.; Ueno, S.; Kakimoto, Y.; Morita, N.
Biochim. Biophys. Acta 1120, 75-80, 1992
A>Title: Saposin-C from bovine spleen; complete amino acid sequence and relation between
A:Reference number: S21770; MUID:92207994; PMID:1554743
A:Accession: S21770
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-80 <SAN>
C:Superfamily: saposin; saposin repeat homology

Query Match 12.4% Score 343; DB 2; Length 80;
Best Local Similarity 79.7%; Pred. No. 4.1e-16;
Matches 63; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 305 SDVCEVCEFLVKEVTKLIDNNKTEKIIDLPDKMSKLPKSLSECEVDTYSSILIS 364
DB 1 ADIYQVEFVAVKAYKIDNNKTEELHMDKVCSTLPSTLAEQCEVDTYSSILIS 60
QY 365 ILLEEVSPELVCSMLTCS 383
DB 61 ILLEEVSPELVCSMLTCS 79

RESULT 8

LNHUB
pulmonary surfactant protein B precursor [validated] - human
N:Alternate names: pulmonary surfactant proteolipid SP-B; pulmonary surfactant-associated
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 08-Dec-2000
C:Accession: A31361; A28461; A27592; J00162; S21134
R:Pilot-Matias, T.J.; Kleiter, S.E.; Fox, J.L.; Kropp, K.; Glaeser, S.W.; Whitsett, J.A.
DNA 8, 75-86, 1989
A>Title: Structure and organization of the gene encoding human pulmonary surfactant protein
A:Reference number: A31361; MUID:89170128; PMID:2924687
A:Accession: A31361
A:Molecule type: DNA
A:Residues: 1-381 <PIB>
A:Cross-references: GB:M24461
A>Note: the codon given for residue 131 (ATT) is inconsistent with the authors' translation
R:Jacobs, K.A.; Phelps, D.S.; Steindrink, R.; Fisch, J.; Kriz, R.; Miltsock, L.; Dougherty, J.
J. Biol. Chem. 262, 9808-9811, 1987
A>Title: Isolation of a cDNA clone encoding a high molecular weight precursor to a 6-kDa
A:Reference number: A28461; MUID:87250653; PMID:3597440
A:Accession: A28461

A:Molecule type: mRNA
A:Residues: 1-227, 'A', 229-381 <JAC>
A:Cross-references: GB:J02761; NID:g190673; PIDN:AAA60212.1; PID:g190674
A>Note: part of this sequence, including the amino end of the mature protein, was confirmed
R:Glaeser, S.W.; Korfhaugen, T.R.; Weaver, T.; Pilot-Matias, T.; Fox, J.L.; Whitsett, J.A.
Proc. Natl. Acad. Sci. U.S.A. 84, 4007-4011, 1987
A>Title: cDNA and deduced amino acid sequence of human pulmonary surfactant-associated protein
A:Reference number: A27794; MUID:87231940; PMID:3035561
A:Accession: A27794
A:Molecule type: mRNA
A:Residues: 'EPF', 99-117, 'L', 319-381 <GLA>
A:Cross-references: GB:M16764; NID:g338410; PIDN:AAA88099.1; PID:g338411

A>Note: 131-Ile was also found
 A>Note: part of this sequence, including the amino end of the mature protein, was confix
 R.Revak, S.D.; Merritt, T.A.; Degryse, E.; Stefani, L.; Courtney, M.; Hallman, M.; Coch
 J. Clin. Invest. 81, 826-833, 1988
 A>Title: Use of human surfactant low molecular weight apoproteins in the reconstitution
 A:Reference number: A27592; MUID:88139786; PMID:3343343
 A:Accession: A27592
 A:Molecule type: mRNA
 A:Residues: 139-177, 'V', 179-227, 'A', 228-381 <REV>
 A:Cross-references: GB:M19097
 A>Note: part of this sequence, including the amino end of the mature protein, was confix
 R.Mizumoto, M.; Adachi, H.
 Sapporo Igaku Zasshi 56, 731-742, 1987
 A>Title: Primary structure of a hydrophobic 6kDa apoprotein (SP6) of human pulmonary su
 A:Reference number: J00162
 A:Accession: J00162
 A:Molecule type: protein
 A:Residues: 201-207, 'X', 209-210, 'X', 212-227, 'A', 229-234, 'X', 236-245, 'X', 247, 'L', 249-253,
 R.Johansson, J.; Joernvall, H.; Curstedt, T.
 FEBS Lett. 301, 165-167, 1992
 A>Title: Human surfactant polypeptide SP-B. Disulfide bridges, C-terminal end, and pep
 A:Reference number: S21134; MUID:92233937; PMID:1568474
 A:Accession: S21134
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 201-227, 'T', 229-279 <JOH>
 A>Note: 228-Ala was also found
 C/Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lowers t
 C/Genetics:
 A:Gene: GDB:SFTPB; SFTPB; SP-B
 A:Cross-references: GDB:120374; OMIM:178640
 A:Map position: 2p12-2p11.2
 A:Intons: 22/1; 65/3; 89/3; 131/3; 194/3; 224/3; 286/1; 334/3; 361/3
 C/Superfamily: pulmonary surfactant protein B; saposin repeat homology
 C/Keywords: alveolar proteinosis; gaseous exchange; glycoprotein; lipoprotein; lung; pul
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-200/Domain: propeptide #status predicted <PRO>
 F:61-153/Domain: saposin repeat homology <SAP1>
 F:200-287/Domain: saposin repeat homology <SAP2>
 F:201-279/Domain: pulmonary surfactant protein B, 9K form #status predicted <SP9>
 F:201-256/Domain: pulmonary surfactant protein B, 6K form #status experimental <SP6>
 F:291-376/Domain: saposin repeat homology <SAP3>
 F:69-143, 72-137, 100-112, 299-366, 302-360, 335-335/Disulfide bonds: #status predicted
 F:129, 311/Binding site: carbohydrate (asn) (covalent) #status predicted
 F:208-277, 211-271, 235-246/Disulfide bonds: #status experimental
 F:248/Disulfide bonds: interchain #status experimental

Query Match 12.3%; Score 339.5; DB 1; Length 381;
 Best Local Similarity 21.4%; Pred. No. 4, 9e-15;
 Matches 109; Conservative 71; Mismatches 169; Indels 161; Gaps 16;

QY 1 LASLIGALAPVGLKCTGSAVWCQNVKTASDCGAVKHCIQTWNKPTVLSLPDICKV 60
 DB 14 LPTLGGPETAAMWTTSSLAACQGFPCWCSLEQALQCRALGHCIQEVWGHGADDL-CQEC 72
 QY 61 KDVTAAQDMKDNATEEELIVLEKTCDMLEPKNMSASCKEIVDSYLPYLIDIKGMSR 120
 DB 73 EDVHIILNKAKKAEIPLQDTMKRKFLEQECNVLPKLMLPQCQVLDVFPVIVDFQWQTD 132
 QY 121 RPEGVCSALNLCESLQKHLAEINHQKQLESNKIPELDMTEVVAFFMANIPLLLYPQDGR 180
 DB 133 SNG-ICMHLGLCKSRQ----- 147
 QY 181 SKPOKNDGVCDQCIGWVDIQTAVRTNSTFVQALVEHKECDRLGPGMADICKNYIS 240
 DB 148 --PEPEGE-----PGMSD----- 158
 QY 241 QYSLEIATQMNMHQPKEICALVGFCDCEVKEMPMQTLVPAKVASKNYIPALEVEPIKKEH 300
 DB 159 -----PLPK-----PLRDLPLPDLKLVLPLP- GAIQARP 189
 QY 301 VPAKSDV-----YCEVCEFLVKEVTKLIDNKTEKEIILDAFDMCKSKLPKSLSEE 350

DB 190 GPHDTLSDEQFPPIPLPFCMLCRALLIRIQAMIRK-----ALRYAVQVRVPLVAGI 245
 QY 351 COEVVDYTGSSILSLILEVSPELVCSMLHICS-----GTRLPALTVAHTQPKDGFCEV 405
 DB 246 CQCLAERSVILLDTLGRMLPOLVCRVLVRCSDSDSAGPRSP---TGEWLPDRSE-CHL 301
 QY 406 CKLVGYIDRNLEKNTKQEFILALBKGC--SFLLPDYQKOCQOFVAFVEFVLIILEV 463
 DB 302 CMSTV-----TQANSSQALPQMLQACGSMWD---REKCKQFVEQHTPQLTLVPRG 353
 QY 464 MDPSEVCLTKGACPSAHKPLIGTEKTIWGP 493
 DB 354 WDAHTTCQALGVCCGTMSPL-----QCIHSP 379

RESULT 9

A29072
 pulmonary surfactant protein SP 18 precursor - dog (fragment)
 C/Species: Canis lupus familiaris (dog)
 C/Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 20-Aug-1999
 C/Accession: B29072; A29072
 R.Hawgood, S.; Benson, B.J.; Schilling, J.; Damm, D.; Clements, J.A.; White, R.T.
 Proc. Natl. Acad. Sci. U.S.A. 84, 66-70, 1987
 A>Title: Nucleotide and amino acid sequences of pulmonary surfactant protein SP 18 and ev
 A:Reference number: A29072; MUID:87092398; PMID:3467361
 A:Accession: B29072
 A:Molecule type: mRNA
 A:Residues: 1-363 <HAM>
 A:Cross-references: GB:M15170; NID:g164077; PIDN:AAA0893.1; PID:g164078
 A:Accession: A29072
 A:Molecule type: protein
 A:Residues: 182-210 <HA2>
 C/Superfamily: pulmonary surfactant protein B; saposin repeat homology
 F:1-14/Domain: signal sequence #status predicted <SIG>
 F:15-180/Domain: propeptide #status predicted <PRO>
 F:54-146/Domain: saposin repeat homology <SAP1>
 F:180-267/Domain: saposin repeat homology <SAP2>
 F:181-363/Domain: pulmonary surfactant protein SP 18 #status experimental
 F:273-358/Domain: saposin repeat homology <SAP3>

Query Match 11.1%; Score 305; DB 2; Length 363;
 Best Local Similarity 21.6%; Pred. No. 8, 7e-13;
 Matches 108; Conservative 75; Mismatches 159; Indels 158; Gaps 19;

QY 5 LGA--LAGPVGLKCTGSAVWCQNVKTASDCGAVKHCIQTWNKPTVLSLPDICKV 63
 DB 13 LGAADMSAPSLA---CARGPAFWCSLEQALQCRALGHCIQEVWGHGADDL-CQECODI 68
 QY 64 VTAAGDMKDNATEEELIVLEKTCDMLEPKNMSASCKEIVDSYLPYLIDIKGMSRPG 123
 DB 69 VAILTKRKKEAIFQDMVAKFLEHEDCVLPKLTLPPQCHNMIGTFPVVVDVFOQIN-PK 127
 QY 124 EYCSALNLCESLQKHLAEINHQKQLESNKIPELDMTEVVAFFMANIPLLLYPQDGRSKP 183
 DB 128 IICKHGLCC--KGLPEPEGESESLPDLKLILPEL----- 162
 QY 184 QPKDNDGVCDQCIGWVDIQTAVRTNSTFVQALVEHKECDRLGPGMADICKNYISQYS 243
 DB 163 -----PGALQVTSPTDLS 177
 QY 244 EIALQMMHMQPKEICALVGFCDCEVKEMPMQTLVPAKVASKNYIPALEVEPIKKEHVA 303
 DB 178 E-----QQLP-----LP----- 186
 QY 304 KSDVYCEVCEFLVKEVTKLIDNKTEKEIID-AFDKCSGLPSLSBECQEVVDYTGSSI 362
 DB 187 ---YCMLCRTLLIRIQAML-----PKGVLAIVTVQGVCHVPLVVGAGICQCLAERYTVLL 237
 QY 363 LSLILEVSPELVCSMLHICS-----GTRLPALTVAHTQPKDGFCEVCKLVGYIDRN 417
 DB 238 LDALLGRMLPOLVGVGLVLRCSHSDSAGPALASPSE-MSFQSEK-CQLCMFVLT-----TQ 290

Qy 418 EKNTKQEIILAEKGC--SFLPDPYOKCDQFAVEYEVILIEIVWMDPSFVCLTKIGA 475
Db 291 AGNHEQATTPQAIIRACLSWMD---RQKCEQFVGHQHRLOTLASGGRDATTQALGA 347
Qy 476 CPSAHKPLGTREKCIWGPSY 495
Db 348 CRTFSPL---QCIIHPHF 363

RESULT 10

146531
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 20-Aug-1999
C/Accession: 146531
R:Margana, R.K.; Boggaram, V.
Am. J. Physiol. 268, L481-L490, 1995
A:Title: Transcription and mRNA stability regulate developmental and hormonal expression
A:Reference number: 146531, MUID:95208794, PMID:7900830
A:Accession: 146531
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-369 <MAN>
A:Cross-references: EMBL:U17106, NID:9642487, PIDN:AAA67934.1, PID:9642488
C:Genetics:
A:Gene: SP-B
C:Superfamily: pulmonary surfactant protein B; saposin repeat homology
F:61-153/Domain: saposin repeat homology <SAP1>

Query Match 10.6%; Score 293; DB 2; Length 369;
Best Local Similarity 21.6%; Pred. No. 5-5e-12;
Matches 110; Conservative 65; Mismatches 165; Indels 170; Gaps 19;

Qy 4 ILGAALAPVVLGL-----KECTRGSAAWQCNVKTASDCGAVHCLQTQWKNKPTVSLPCD 58
Db 12 LLLPPLCGPATAVWATSPACAGQGFQCSLEQLCKALGHCLQEVGHGAGADL-CQ 70
Qy 59 ICKDVVTAAGMLKDNATEEBILVYLEKTCMLPRPNNSASCKEIVDSYLPVILDIIGE 118
Db 71 ECQDVLNLTQWTKKAIPODITRKFLHEBCDVLPLKLLVPOCHVLDVYFPPLTITYF 130
Qy 119 MSRPVEVSALNLCESLQKHLAELNHQKLESNKIPEDMTEVAVPAMNIPILLYPDG 178
Db 131 INAKA-ICQHLGLCQ-----PQSEPPPLD-----PLPDKLVLPFL---GA 167
Qy 179 PRSKRPQPKNDQVCDICIQWTDIQTAVRNTSTFQVALVHVKESCDRLGPMADICKY 238
Db 168 LPAPK-----GPHTDLS--- 180
Qy 239 ISOYSEINAIQMMHMQKEICALVGFCEVKEPMQOTLVPAKVASKNVIIPALELVEPIKK 298
Db 181 -----AQRFPPL-----PL-- 190
Qy 299 HEVPAKSDVCEVCEFLVKEVTKLIDNNKTEKEILD-AFDKMCSTLPSLSSECOEVDVT 357
Db 191 -----CMCKCTLAKRIQAMT-----PKGVLAAMAVAGCHVPLVVGICQRLAKR 235
Qy 358 YGSSILSLLEBVSPELVCSMLHLC-----GTRLPALTVAHTQ--PKDGFCEVCKLV 410
Db 236 YTVILLLEVLGHVLPOLVCGVLRCCSVDSIGQVPTPEALPGEWLPDPPE-CRLCMVTV 294
Qy 411 GYLDRLNEKNTKQEIILAA-----LEKGSFLPDPYOKCDQFAVEYEVILIEIVWMD 465
Db 295 TQA-NNISEQTRPQAVVHACLSQDCK-----OECRQFVELHTPQLLSLRGMD 343
Qy 466 PSFVCLKIGACPSAHKPLGTREKCIWGPSY 495
Db 344 ARAIQALGACVATISPL-----QCIGSPHF 369

RESULT 11
A32026
glucosylceramide beta-glucosidase activator protein SAP-2 - guinea pig

C:Species: Cavia porcellus (guinea pig)
C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 17-May-1996
C/Accession: A32026
R:Sano, A.; Radin, N.S.; Johnson, L.L.; Tarr, G.E.
J. Biol. Chem. 263, 19597-19601, 1988
A:Title: The activator protein for glucosylceramide beta-glucosidase from guinea pig liv
A:Reference number: A32026, MUID:89066787, PMID:3198642
A:Accession: A32026
A:Molecule type: protein
A:Residues: 1-81 <SAN>
C:Superfamily: saposin; saposin repeat homology
F:1-81/Domain: saposin repeat homology <SAP>

Query Match 9.5%; Score 261; DB 2; Length 81;
Best Local Similarity 61.8%; Pred. No. 1.1e-10;
Matches 47; Conservative 18; Mismatches 11; Indels 0; Gaps 0;

Qy 307 VYCEVCEFLVKEVTKLIDNNKTEKEILDAPFMCKSLPKSISECOEVDVYGSISIL 366
Db 3 VTCACACEVVKKWMELIDNNKTEKEIHALDSVCALPESVSEVCEVVDYIGDSIVALL 62
Qy 367 LEEVSPPELVCSMLHLC 382
Db 63 LQEMSPPELVCSMLHLC 78

RESULT 12

INRRB

pulmonary surfactant protein B precursor - rabbit
N:Alternate names: pulmonary surfactant-associated protein-B
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 22-Jun-1999
C/Accession: A32421
R:Xu, J.; Richardson, C.; Ford, C.; Spencer, T.; Li-Juan, Y.; Mackie, G.; Hammond, G.; P
Biochem. Biophys. Res. Commun. 160, 325-332, 1989
A:Title: Isolation and characterization of the cDNA for pulmonary surfactant-associated
A:Reference number: A32421, MUID:89228033, PMID:2469419
A:Accession: A32421
A:Molecule type: mRNA
A:Residues: 1-370 <XU>
A:Cross-references: GB:M24901, NID:9165707, PIDN:AAA31466.1, PID:9165708
A>Note: the authors translated the codon CCG for residue 184 as Arg and CAG for residue
C:Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lowers t
C:Superfamily: pulmonary surfactant protein B; saposin repeat homology
C:Keywords: alveolar proteinosis; gaseous exchange; glycoprotein; lipoprotein; lung; pul
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-184/Domain: propeptide #status predicted <PRO>
F:62-154/Domain: saposin repeat homology <SAP1>
F:184-271/Domain: saposin repeat homology <SAP2>
F:185-263/Product: pulmonary surfactant protein B, 9K form #status predicted <SP9>
F:185-240/Product: pulmonary surfactant protein B, 6K form #status predicted <SP6>
F:280-365/Domain: saposin repeat homology <SAP3>
F:300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.1%; Score 250; DB 1; Length 370;
Best Local Similarity 20.7%; Pred. No. 3.8e-09;
Matches 108; Conservative 63; Mismatches 162; Indels 190; Gaps 21;

Qy 1 LASLGAALAPVVLGL-----KECTRGSAAWQCNVKTASDCGAVHCLQTQWKNKPTVSL 55
Db 10 LLLPPLCGPATAVWATSPACAGQGFQCSLEQLCKALGHCLQEVGHGAGADL 69
Qy 56 PCDICQDVVTAAGMLKDNATEEBILVYLEKTCMLPRPNNSASCKEIVDSYLPVILDI 115
Db 70 -CQECQDVLNLTQWTKKAIPODITRKFLHEBCDVLPLKLLVPOCHVLDVYFPPLTITYF 128
Qy 116 KGEVSPPEVSALNLCESLQKHLAELNHQKLESNKIPEDMTEVAVPAMNIPILLYP 175
Db 129 QSQINAKA-ICQHLGLCQ-----PQSEPPPLD-----PLPDKLVLPFL--- 166
Qy 176 QDGPSPKQPKNDQVCDICIQWTDIQTAVRNTSTFQVALVHVKESCDRLGPMADIC 235
Db 167 -GALPAKRGF----- 175

QY 236 KNYISQYSEIALQMMHMQPKICALVGFCDYKEMQNTLVPAKVASKNVIALBELVER 295
Db 176 -----HQ-----DLQAQFPPIPLP-----190
QY 296 IKKHEVPAKSDYVCEVCEFLVKEVTKLIDNNKTEKEIID-AFDMGSKLPKSLSECOY 354
Db 191 -----LCMLCRTLLKRIQAMI-----PKGVLAAVAVOVCHVAVELVGGICQCL 233
QY 355 VDTYGSSILSTILLEEVSPELVCSMLHCS-----GTRLPALVAVTQ--PKDGGFCEVCK 407
Db 234 AERTVTLLEVLGHVLPQVLCGLVLRCSVDSIGVPTLEALPGEMLPQDDE-CRLCM 292
QY 408 KLVGYIDRLNLEKSTKOEIIAA-----LEKGCSTLPDPYQKQDQFAVEEPLIILIVE 462
Db 293 SVTTQA-RNISEQRPAQVAVHACISSQLDK-----QCEQCFVAAHAP-----333
QY 463 VMDPSFVCLKIGACPSAKPLLGT-----KCIWGPSY 495
Db 334 AAEPAVQGL---GCP---RNLPGEGRVAVATLSPLQICQSPHF 370

RESULT 13

T46069
hypothetical protein T18N14.110 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T46069
R:Deleeny, M.; Berger, C.; Cooke, R.; Grellet, F.; Laurie, M.; Mewes, H.W.; Lemcke, K.;
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23013
A:Accession: T46069
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213
A:Cross-references: EMBL:AL132968
A:Experimental source: cultivar Columbia; BAC clone T18N14
C:Genetics:
A:Map position: 3
A:introns: 31/1; 80/3; 146/3; 166/3
A>Note: T18N14.110

Query Match 7.7%; Score 213; DB 2; Length 213;
Best Local Similarity 29.7%; Pred. No. 5.1e-07;
Matches 57; Conservative 35; Mismatches 90; Indels 10; Gaps 6;

QY 286 VIPALELVEPIKHEVPAKSDYVCEVCEFLVKEVTKLIDNNKTEKEIIDAFDMGSKLPK 345
Db 18 VSDARSFVDSITSEKVSNKEDV-CTLCEEVTDALSYLEKVVQAEIIBDLHRCQGL-R 75
QY 346 SLSECCQEVDTYGGSSILSTILLEEVSPELVCSMLHCSGTRLPALVAVTQPKDGGFCEV 405
Db 76 GYSQGCISLVVY-VPLFQLQESQPHYFCRKNLQG-KVVALVEARQDS---CGV 128
QY 406 CKLVGYIDRLNLEKSTKOEIIAALEKGCSTLPDPYQKQDQFAVEEPLIILIVEVD 465
Db 129 CHRVSSELLIKQDPDQDLDIVELLIKGCKSLKN-YEKCKTIVFEIGPILVNAEEFLV 187
QY 466 PSFVCLKIGACP 477
Db 188 KNDVCTILRACP 199

RESULT 14

T48201
hypothetical protein T20L15.70 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T48201
R:Bevan, M.; Peters, S.A.; van Steveren, M.; Dirks, W.; Stiekema, W.; Bancroft, I.; Mew
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24488
A:Accession: T48201

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-217 <BEV>
A:Cross-references: EMBL:AL162351
A:Experimental source: cultivar Columbia; BAC clone T20L15
C:Genetics:
A:Map position: 5
A:introns: 30/1; 79/3; 146/3; 166/3
A>Note: T20L15.70

Query Match 7.4%; Score 203.5; DB 2; Length 217;
Best Local Similarity 27.8%; Pred. No. 2.2e-06;
Matches 52; Conservative 40; Mismatches 82; Indels 13; Gaps 7;

QY 292 LVEPIKK-HEVPAKSDYVCEVCEFLVKEVTKLIDNNKTEKEIIDADDMGSKLPKSLSEB 350
Db 25 LLEPFESAHND---DNQVCELCQKVTYTLVDYLDQDNQNELVLAHISCQIP-PLKIQ 79
QY 351 QCEVDTYGGSSILSTILLEEVSPELVCSMLHCSGTRLPALVAVTQPKDGGFCEVCKLY 410
Db 80 CLSMVDHY-TQLPFTQVSTIKSDQICKRLNLQAV-TPAFASQVHQ---GNCEACRETV 133
QY 411 GYLDRLNLEKSTKOEIIAALEKGCSTLPDPYQKQDQFAVEEPLIILIVEVMDPSFVC 470
Db 134 SEVTKLKDPEYTKLIRLLLECKSL-NNYQDKCKWFEYGPMLTDLQFLERKQVC 192
QY 471 LKIGACP 477
Db 193 TIHVCIP 199

RESULT 15

T15677
hypothetical protein C28C12.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15677
R:Miller, N.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C28C12.
A:Reference number: Z18387
A:Accession: T15677
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-402 <MULT>
A:Cross-references: EMBL:U40797; NID:g1065916; PID:g1065921; PIDN:AA837548.1; GSPDB:GN001
A:Experimental source: strain Bristol N2; clone C28C12
C:Genetics:
A:gene: CESP:C28C12.5
A:Map position: 4
A:introns: 29/3; 82/3; 124/3; 151/3; 258/3; 318/1; 372/3; 400/3

Query Match 6.4%; Score 176.5; DB 2; Length 402;
Best Local Similarity 20.3%; Pred. No. 0.00029;
Matches 75; Conservative 58; Mismatches 156; Indels 81; Gaps 15;

QY 178 GPRSKPQKNDGVCOQDICIQVATVTRNSTFVQALVEHVEEDRQGPMAQICN 237
Db 16 GAQSAFSP-----CECKSKVQNFIDASKDRMAQLKVS-LSMLC--VGTSHQSDCK 66
QY 238 YISQYSEIALQMMHMQPKICALVGFCDYKEMQNTLVPAKVASKNVIALBELVERP 296
Db 67 TLIDKIDFLAYKLAAYLADTSVAVCSKIQMGESQSPILARLAMLVLKSEIVA---NDNI 123
QY 297 KKEVPAKSDYVCEVCEFLVKEVTKLIDNNKTEKEIIDADDMGSKLPKSLSECCQEVVD 356
Db 124 MRQEV-----CEQCASTAQIGKLVGDEFTYAVVSTLQRFVCSAGKAAHAGC-----N 172
QY 357 TYGSSILSTILLEEVSPELVCSMLHCSGTRLPALVAVTQPKDGGFCEVCKLVG 411
Db 173 IFVSVIPDLMTWKMDFTKEKLMGCSNMGCSATSKPARARAPKQAS---EMKSM-- 226
QY 412 YLDRLNLEKSTKOEIIAALEKGCSTLPDPYQKQ----- 444

```
Db      227 ----GMVXTSNGEELMSCFE--CTLSADALLQEFIDXRGTADDIQTVACNMGVANWTDG 280
QY      445 CDQFVAEYEPVLIETLVEWMDPSFVCLKIGACPSAHKPLG--TEKCIWGPSYWCNTE 501
Db      281 CNDFVHMWMTSTVLFLTYNQFDGRGICTVMHSCCKENALVEWAMSEKVMLG---CEN-- 334
QY      502 TAAQCNAVEH 511
Db      335 ----CKAVEH 340
```

Search completed: May 5, 2004, 13:34:35
Job time : 22.661 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:16:48 ; Search time 12.7901 Seconds
(without alignments)
2108.841 Million cell updates/sec

Title: US-09-743-684a-1_COPY_7_524
Perfect score: 2759
Sequence: 1 LASILGALAGPVLGKRECT.....NTEPAQCNAYEHCKRHVWN 518

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2759	100.0	524	1 SAP_HUMAN	P07602 h proactiva
2	2405.5	87.2	525	1 SAP_BOVIN	P26779 b proactiva
3	1974	71.5	554	1 SAP_RAT	P10960 rattus novy
4	1916.5	65.5	557	1 SAP_MOUSE	P61207 mus musculu
5	1676.5	60.8	518	1 SAP_CHICK	O13035 gallus gall
6	393	14.2	80	1 SAP_PIG	P81405 sus scrofa
7	378	13.7	377	1 PSPB_MOUSE	P50405 mus musculu
8	372	13.5	376	1 PSPB_RAT	P22355 rattus novy
9	340.5	12.3	381	1 PSPB_HUMAN	P17129 canis famli
10	305	11.1	363	1 PSPB_CANFA	P15285 oryctolagus
11	290	10.5	370	1 SAP_CAVPO	P20097 cavia porce
12	261	9.5	81	1 SAP_RABIT	P42210 hordeum vul
13	173.5	6.3	513	1 ASPR_CUCPE	P42082 cynara card
14	158.5	5.7	508	1 ASPR_HORVU	P42211 oryza sativ
15	154.5	5.6	473	1 CYPL_CYNCA	P61343 mus musculu
16	154	5.6	496	1 ASPR_ORYSA	O42456 oryza sativ
17	143.5	5.2	1175	1 GLG1_MOUSE	O62338 rattus novy
18	142.5	5.2	509	1 APR1_ORYSA	O92196 homo sapien
19	138.5	5.0	1171	1 GLG1_RAT	O92196 homo sapien
20	137.5	5.0	1179	1 GLG1_HUMAN	O92196 homo sapien
21	137.5	5.0	1160	1 GLG1_CRIGR	O92196 homo sapien
22	127.5	4.6	1142	1 GLG1_CHICK	O83875 enterococu
23	120	4.3	578	1 EZRA_CHICK	O83875 enterococu
24	119	4.3	975	1 SECB_MOUSE	O35382 mus musculu
25	118	4.3	857	1 I2C1_HUMAN	O94118 mus musculu
26	117	4.2	857	1 I2C1_MOUSE	O94118 mus musculu
27	116.5	4.2	1216	1 P1B1_MOUSE	O94118 mus musculu
28	116	4.2	975	1 SECB_RAT	O62824 rattus novy
29	114.5	4.2	971	1 RECK_HUMAN	O95860 homo sapien
30	114	4.1	984	1 NOR_HUMAN	P16320 dictyosilla
31	113.5	4.1	974	1 SECB_HUMAN	O96655 homo sapien
32	113	4.1	418	1 TEK1_HUMAN	O96654 homo sapien
33	113	4.1	1102	1 MYSC_CHICK	P29616 gallus gall

34	113	4.1	2869	1 RBP1_PLAVB	Q00798 plasmodium
35	112	4.1	1184	1 FBI2_HUMAN	P98095 homo sapien
36	112	4.1	1216	1 P1B1_BOVIN	P10894 bos taurus
37	112	4.1	1216	1 P1B1_RAT	P10687 rattus novy
38	111.5	4.0	8545	1 ANCI1_CAEEL	O944m4 caenorhabdi
39	110.5	4.0	3911	1 AKR9_HUMAN	O99996 h-a-kinase
40	110	4.0	476	1 VTDB_RABIT	P53789 oryctolagus
41	110	4.0	1216	1 P1B1_HUMAN	O94118 mus musculu
42	109	4.0	861	1 I2C4_HUMAN	O94118 mus musculu
43	109	4.0	1557	1 DVAI1_DICVI	O94118 mus musculu
44	108.5	3.9	3674	1 SPGR_HUMAN	O94118 mus musculu
45	108.5	3.9	3678	1 DMD_MOUSE	P11531 mus musculu

ALIGNMENTS

RESULT 1
SAP_HUMAN STANDARD: PRT: 524 AA
AC P07602; P07292; P15793; P78538; P78541; P78546; P78547; P78558;
AC O92739; O92740; O92741; O92742;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Proactivator polypeptide precursor [contains: Saposin A (Protein A);
DE Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside
DE sulfate activator) (CSact) (Dispersin) (Sulfatide/GM1 activator);
DE Saposin C (Co-beta-glucosidase) (AI activator) (Glucosylceramidase
DE activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D
DE (Protein C) (Component C)].
GN PSAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=Liver;
RC MEDLINE=90129043; PubMed=2515150;
RA Roman E.G., Grabowski G.A.;
RT "Molecular cloning of a human co-beta-glucosidase cDNA: evidence that
RT four sphingolipid hydrolase activator proteins are encoded by single
RT genes in humans and rats.";
RL Genomics 5:486-492 (1989).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89255151; PubMed=2498298;
RA Nakano T., Sandhoff K., Struemper J., Christomanou H., Suzuki K.;
RT "Structure of full-length cDNA coding for sulfatide activator, a
RT co-beta-glucosidase and two other homologous proteins: two alternate
RT forms of the sulfatide activator.";
RL J. Biochem. 105:152-154 (1989).
[3]
RP SEQUENCE FROM N.A. (ISORF SAP-MU-0).
RC TISSUE=Brain, Eye, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Struenser R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marnett K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshimori S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Boesek S.A., McKernan K.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Faney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Bialesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalske U., Smallos D.E.,

- RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 59-125 AND 304-513 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=91192146; PubMed=2013321;
RA Holtschmidt H., Sandhoff K., Fuertst W., Kwon H.Y., Schnabel D.,
RT Suzuki K.;
RL "The organization of the gene for the human cerebroside sulfate
RT activator protein.";
RL FEBS Lett. 280:267-270(1991).
RN [5]
RP SEQUENCE OF 164-524 FROM N.A.
RX MEDLINE=88068647; PubMed=2825202;
RA Dewi N.N., Wenger D.A., O'Brien J.S.;
RT "Nucleotide sequence of cloned cDNA for human sphingolipid activator
RT protein 1 precursor.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8652-8656(1987).
RN [6]
RP PARTIAL SEQUENCE OF 60-142.
RX MEDLINE=89240739; PubMed=2717620;
RA Morimoto S., Martin B.M., Yamamoto Y., Kretz K.A., O'Brien J.S.,
RT Kishimoto Y.;
RL "Saposin A: second cerebroside activator protein.";
RN Proc. Natl. Acad. Sci. U.S.A. 86:3389-3393(1989).
RP [7]
RP SEQUENCE OF 195-263 FROM N.A.
RX MEDLINE=86130593; PubMed=2668718;
RA Dewi N.N., Wenger D.A., Fujibayashi S., Donoviel M., Esch F.,
RT Hill F., O'Brien J.S.;
RL "Molecular cloning of the sphingolipid activator protein-1 (SAP-1),
RT the sulfatide glucosylase activator.";
RN Biochem. Biophys. Res. Commun. 134:989-994(1986).
RP [8]
RP SEQUENCE OF 195-274.
RC TISSUE=Kidney;
RX MEDLINE=91006165; PubMed=2209618;
RA Furst W., Schubert J., Machleidt W., Meyer H.E., Sandhoff K.;
RT "The complete amino-acid sequences of human ganglioside GM2 activator
RT protein and cerebroside sulfate activator protein.";
RN Eur. J. Biochem. 192:709-714(1990).
RP [9]
RP SEQUENCE OF 195-274.
RX MEDLINE=89207118; PubMed=3242555;
RA Kleinschmidt T., Christomanou H., Braunitzer G.;
RT "Complete amino-acid sequence of the naturally occurring A2 activator
RT protein for enzymic sphingomyelin degradation: identity to the
RT sulfatide activator protein (SAP-1).";
RL Biol. Chem. Hoppe-Seyler 369:1361-1365(1988).
RN [10]
RP SEQUENCE OF 311-390.
RX MEDLINE=88163077; PubMed=3442600;
RA Kleinschmidt T., Christomanou H., Braunitzer G.;
RT "Complete amino-acid sequence and carbohydrate content of the
RT naturally occurring glucosylceramide activator protein (A1 activator)
RT absent from a new human Gaucher disease variant.";
RL Biol. Chem. Hoppe-Seyler 368:1571-1578(1987).
RN [11]
RP SEQUENCE OF 407-484.
RX MEDLINE=89000190; PubMed=3048308;
RA Furst W., Machleidt W., Sandhoff K.;
RT "The precursor of sulfatide activator protein is processed to three
RT different proteins.";
RL Biol. Chem. Hoppe-Seyler 369:317-328(1988).
RN [12]
RP PARTIAL SEQUENCE OF 405-484.
RX MEDLINE=89025876; PubMed=2845979;
RA Morimoto S., Martin B.M., Kishimoto Y., O'Brien J.S.;
RT "Saposin D: a sphingomyelinase activator.";
RL Biochem. Biophys. Res. Commun. 156:403-410(1988).
RN [13]
RP SEQUENCE OF 17-26.
RC TISSUE=Milk;
RX MEDLINE=92068206; PubMed=1958198;
RA Kondoh K., Hineno T., Sano A., Kakimoto Y.;
RT "Isolation and characterization of prosaposin from human milk.";
RL Biochem. Biophys. Res. Commun. 181:286-292(1991).
RN [14]
RP PARTIAL SEQUENCE (SAPOSIN B), AND STRUCTURE OF CARBOHYDRATES.
RC TISSUE=urine;
RX MEDLINE=20032116; PubMed=10562467;
RA Fluharty A.L., Lombard C., Louis A., Stevens R.L., Whitelegge J.P.,
RT Waring A.J., To T., Fluharty C.B., Paul K.F.;
RL "Preparation of the cerebroside sulfate activator (CSACT or saposin B)
RT from human urine.";
RN Mol. Genet. Metab. 68:391-403(1999).
RP [15]
RP STRUCTURE OF CARBOHYDRATE ON ASN-215.
RX MEDLINE=21110404; PubMed=11180632;
RA Paul K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,
RT Stevens R.L., Fluharty C.B., Fluharty A.L.;
RL "Structure of the asparagine-linked sugar chains of porcine kidney and
RT human urine cerebroside sulfate activator protein.";
RN J. Mass Spectrom. 35:1416-1424(2000).
RP [16]
RP SAPOSIN D DISULFIDE BONDS.
RX MEDLINE=9937688; PubMed=10406958;
RA Tatti M., Salvioli R., Claffoni F., Pucci P., Andolfo A.,
RT Amoresano A., Vaccaro A.M.;
RL "Structural and membrane-binding properties of saposin D.";
RN Eur. J. Biochem. 263:486-494(1999).
RP [17]
RP SAPOSIN B DISULFIDE BONDS.
RX MEDLINE=2398398; PubMed=12510003;
RA Ahn V.E., Paul K.F., Whitelegge J.P., Higginson J., Fluharty A.L.,
RT Prive G.G.;
RL "Expression, purification, crystallization, and preliminary X-ray
RT analysis of recombinant human saposin B.";
RN Protein Expr. Purif. 27:186-193(2003).
RP [18]
RP MASS SPECTROMETRY.
RC TISSUE=Urine;
RX MEDLINE=99441404; PubMed=10510427;
RA Paul K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,
RT Krutichinsky A.N., Standing K.G., Waring A.J., Stevens R.L.,
RL Fluharty C.B., Fluharty A.L.;
RN "Cerebroside sulfate activator protein (Saposin B): chromatographic
RT and electrospray mass spectrometric properties.";
RL J. Mass Spectrom. 34:1040-1054(1999).
RP [19]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 195-273, AND MUTAGENESIS OF
RP ILE-240.
RX MEDLINE=22406333; PubMed=12518053;
RA Ahn V.E., Paul K.F., Whitelegge J.P., Fluharty A.L., Prive G.G.;
RT "Crystal structure of saposin B reveals a dimeric shell for lipid
RT binding.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:38-43(2003).
RN [20]
RP REVIEW ON MLD VARIANTS.
RX MEDLINE=95170731; PubMed=7866401;
RA Gieselmann V., Zlotogora J., Harris A., Wenger D.A., Morris C.P.;
RT "Molecular genetics of metachromatic leukodystrophy.";
RL Hum. Mutat. 4:233-242(1994).
RN [21]
RP VARIANT MLD ILE-217.
RX MEDLINE=90147748; PubMed=2302219;
RA Rafi M.A., Zhang X.-L., Degala G., Wenger D.A.;
RT "Detection of a point mutation in sphingolipid activator protein-1
RT mRNA in patients with a variant form of metachromatic
RT leukodystrophy.";
RL Biochem. Biophys. Res. Commun. 166:1017-1023(1990).
RN [22]
RP SEQUENCE FROM N.A., AND VARIANT MLD ILE-217.
RX MEDLINE=90207231; PubMed=2320574;

FT DISULFID 231 242 BY SIMILARITY.
 FT DISULFID 316 389 BY SIMILARITY.
 FT DISULFID 319 383 BY SIMILARITY.
 FT DISULFID 347 358 BY SIMILARITY.
 FT DISULFID 410 483 BY SIMILARITY.
 FT DISULFID 413 477 BY SIMILARITY.
 FT DISULFID 441 452 BY SIMILARITY.
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 427 427 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 127 127 H -> R.
 FT VARIANT 127 127 MOCK -> IRIR.
 FT VARIANT 260 263 E -> Q (IN REF. 2).
 FT CONFLICT 317 317 R -> S (IN REF. 1).
 FT CONFLICT 367 367
 SQ SEQUENCE 525 AA; 58120 MW; 293AF0FB9C4F99 CRC64;

Query Match 87.2%; Score 2405.5; DB 1; Length 525;
 Best Local Similarity 84.6%; Pred. No. 2.5e-150;
 Matches 439; Conservative 43; Mismatches 36; Indels 1; Gaps 1;

QY 1 LASTLGAALAGAVLGLKCTRGSAVCQVNTASDCAVAGCLQTVNKPVYKSPIC 60
 DB 7 LASTLGAALASVPLGLRECTRGSAVCQVNTAAGVAGVHCLQTVNKPVYKSPIC 66
 QY 61 KDVTAAAGDMLKDNATEEELVYLEKTCVPLPKPMASCKEIVDSYLPVLDITK 120
 DB 67 KDVTAAAGDMLKDNATEEELVYLEKTCVPLPKPMASCKEIVDSYLPVLDITK 126
 QY 121 RRGVCSALNLCESIQKHLAEINHOKEENKIPELDMTEVAPPMANIPLLVYPOD 180
 DB 127 HGEVCSALNLCESIQKHLAEINHOKEENKIPELDMTEVAPPMANIPLLVYPOD 186
 QY 181 SKPGKRD-NGDVCDICQVNTDIOFVATNSFVQALVEHVKEDRLGPGKADIC 239
 DB 187 SKPGKRD-NGDVCDICQVNTDIOFVATNSFVQALVEHVKEDRLGPGKADIC 246
 QY 240 SOYSSIALIOMMMQPKICALVGFCDVCEKEMQTVLPAYAKVKNVPALEVEPIK 299
 DB 247 NOYSEVALIOMMMQPKICALVGFCDVCEKEMQTVLPAYAKVKNVPALEVEPIK 306
 QY 300 EVPAKSDVCEVCEVCEVCEVCEVCEVCEVCEVCEVCEVCEVCEVCEVCEVCE 359
 DB 307 PAPADADYCEVCEVCEVCEVCEVCEVCEVCEVCEVCEVCEVCEVCEVCEVCE 366
 QY 360 SSILSILEEVSPELVCSMLHLCSGTRLPALVHYVTPKQDGFCEVCKKLVYDRL 419
 DB 367 RSILSILDEASPELVCSMLHLCSGTRLPALVHYVTPKQDGFCEVCKKLVYDRL 426
 QY 420 NSTKQEIILAEKGCSPDPYQKQDGFVAEYEPVLEILVEVMDSPFVCLKIGAC 479
 DB 427 NSTKQEIILAEKGCSPDPYQKQDGFVAEYEPVLEILVEVMDSPFVCLKIGAC 486
 QY 480 HKPILGTEKCIWPSYVCONMETAAQCAVAGHCHRWAN 518
 DB 487 HKPILGTEKCIWPSYVCONMETAAQCAVAGHCHRWAN 525

RESULT 3
 SLP_RAT STANDARD; PRT; 554 AA.
 AC P10560; O62841; O64190;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DE 15-MAR-2004 (Rel. 43, Last annotation update)
 GN PSAP OR SGP1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]

RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC TISSUE=Sertoli cells;
 RA MEDLINE=89000647; PubMed=3048385;
 RX Collard M.W., Sylvester S.R., Tsuruta J.K., Griswold M.D.,
 RT "Biosynthesis and molecular cloning of sulfated glycoprotein 1
 RT 70-kilodalton precursor to sulfatide/GM1 activator".
 RL Biochemistry 27:4557-4564(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA MEDLINE=96128541; PubMed=8573994;
 RX Morales C.R., El-Alfy M., Zhao Q., Igdeura S.A.,
 RT "Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in
 RT Sertoli cells".
 RL Histol. Histopathol. 10:1023-1034(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testicle;
 RA MEDLINE=96175245; PubMed=8601692;
 RX Morales C.R., El-Alfy M., Zhao Q., Igdeura S.A.,
 RT "Expression and tissue distribution of rat sulfated glycoprotein-1
 RT (prosaposin)".
 RL J. Histochem. Cytochem. 44:327-337(1996).
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- SIMILARITY: Contains 2 saposin A-type domains.
 CC -1- SIMILARITY: Contains 4 saposin B-type domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M19936; AAA2136.1; -
 DR EMBL; S81353; AAB36042.2; -
 DR EMBL; S81373; AAB36235.2; -
 DR PIR; A28716; A28716.
 DR InterPro; IPR003119; Sapa.
 DR InterPro; IPR007856; Sapa_1.
 DR InterPro; IPR008138; Sapa_2.
 DR InterPro; IPR008140; Sapa_sub.
 DR InterPro; IPR008373; Saposin.
 DR InterPro; IPR008139; Saposinb.
 DR Pfam; PF02199; SAPA; 2.
 DR Pfam; PF05184; Sapa_1; 3.
 DR Pfam; PF03489; Sapa_2; 4.
 DR PRINTS; PR01797; SAPOSIN.
 DR ProDom; PD001732; Sapa_sub; 3.
 DR SMART; SM00162; SAPA; 2.
 DR SMART; SM00118; SApB; 4.
 KW Signal; Glycoprotein; Repeat.
 FT SIGNAL 1 16
 FT CHAIN 17 554 SULFATED GLYCOPROTEIN 1.
 FT DOMAIN 21 54 SAPOSIN-LIKE TYPE A 1.
 FT DOMAIN 61 138 SAPOSIN-LIKE TYPE B 1.
 FT DOMAIN 193 274 SAPOSIN-LIKE TYPE B 2.
 FT DOMAIN 310 391 SAPOSIN-LIKE TYPE B 3.
 FT DOMAIN 435 516 SAPOSIN-LIKE TYPE B 4.
 FT DOMAIN 521 554 SAPOSIN-LIKE TYPE A 2.
 FT DISULFID 63 138 BY SIMILARITY.
 FT DISULFID 66 132 BY SIMILARITY.
 FT DISULFID 94 106 BY SIMILARITY.
 FT DISULFID 197 270 BY SIMILARITY.
 FT DISULFID 200 264 BY SIMILARITY.
 FT DISULFID 229 240 BY SIMILARITY.
 FT DISULFID 314 387 BY SIMILARITY.
 FT DISULFID 317 381 BY SIMILARITY.
 FT DISULFID 345 356 BY SIMILARITY.
 FT DISULFID 439 512 BY SIMILARITY.
 FT DISULFID 442 506 BY SIMILARITY.

```

FT DISULFID 470 481 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 115 115 P -> L (IN REF. 2 AND 3).
FT CONFLICT 299 299 D -> E (IN REF. 2).
FT CONFLICT 462 462 I -> V (IN REF. 3).
FT CONFLICT 527 527 W -> R (IN REF. 3).
FT CONFLICT 536 536 S -> M (IN REF. 3).
SQ SEQUENCE 554 AA; 61123 MW; DFE3FA3A0520C6B CRC64;

Query Match 71.5%; Score 1974; DB 1; Length 554;
Best Local Similarity 66.1%; Pred. No. 4,8e-122;
Matches 363; Conservative 77; Mismatches 77; Indels 32; Gaps 3;

QY 1 LASLIGALAGVIGLKECTRGSAVWQNVKTASDCGAVKHCLOTVWKNKPKVSLPCDIC 60
DB 7 LASLIVTALTSVPQPKICSGGSAVVCRCVKTAVDCRAVKHCQGVWWSKPTAKSLPCDIC 66
QY 61 KDVTAAAGDMKDNATEEELIYLEKTCDDMLPKPMMSACEIYDSYLPVILDIKIGMS 120
DB 67 KTVVTEANLNDKNTTEEELIYLEKTCAMHDSLSASCKEVDSYLPVILDMKIGMS 126
QY 121 RPEVCALNLCESLQKHLAEIHNOKLESNKIPILDMTEVVAAPPANIPILLYPODGR 180
DB 127 NPEVCALNLCESLQKHLAEIHNOKLESNKIPILDMTEVVAAPPANIPILLYPODGR 185
QY 181 SKPOKNDGVQCQDCIOWVTIQTAVRTNSTFVQALVEHVKEECRLGPGMADICKNYIS 240
DB 186 SOPKRNEDVQCQDCMKLVTDIQTAVRTNSTFVQALVEHVKEECRLGPGVADICKNYVD 245
QY 241 QYSEVALIOMMMHMOKEICALVGFCEVKEPMQTLVPAKASKAVIPALVEIKKHE 300
DB 246 QYSEVALIOMMMHMOKEICVWVGFCEVKEPMQTLVPAKASKAVIPALVEIKKHE 305
QY 301 VPAKSDVCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCLPKSLSECOEVDYTGSS 360
DB 306 IQQNVITFCQVQALVWKLSEILINNATEELIKGLSKRCSLLPAPASTKQCEVLTGEP 365
QY 361 SILSILLESVPELVCSMLHLCGSG-----TRLP-----A 389
DB 366 SILDVLMEHVNPNFLCGYISLCSANPNLVGLEPPAAIVSALPREPAPKQPEBPQKSA 425
QY 426 LRAHVPQKNGGFCVCKKLVYLEHNEKXSTKEIILAALKGCSFLPDPYQXCDERFV 485
DB 450 AVEEVLLEILVEVMDPSFVCLKIGACPSAHKPLGTGTEKICMGPSYWCNTEATAQCNV 509
QY 486 AVEEVLLEILVEVMDPSFVCKSIGVCSAYKLLGTGTEKICMGPSYWCNSETARCNV 545
DB 510 EHCKRHVNN 518
QY 546 DHCKRHVNN 554

```

```

RX MEDLINE=92272718; PubMed=1590788;
RA Tsuda M., Sakiyama T., Endo H., Kitagawa T.;
RT "The primary structure of mouse saposin A";
RL Biochem. Biophys. Res. Commun. 184:1266-1272(1992).
(2)
RP SEQUENCE FROM N.A.
RX MEDLINE=94272317; PubMed=8003952;
RA Sprecher-Levy H., Orr-Urtreger A., Lonal P., Horowitz M.;
RT "Murine prosaposin: expression in the reproductive system of a gene
implicated in human genetic disease";
RL Cell. Mol. Biol. 40:233-233(1994).
(3)
RP SEQUENCE FROM N.A.
RX MEDLINE=96084310; PubMed=8565332;
RA Cao Q.P., Crain W.R.;
RT "Expression of SGP-1 mRNA in preimplantation mouse embryos";
RL Dev. Genet. 17:263-271(1995).
(4)
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Zhao Q.Q., Hay N.N., Morales C.R.;
RT Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: Contains 2 saposin A-type domains.
CC -1- SIMILARITY: Contains 4 saposin B-type domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; S36200; AAB22175.1; -
CC EMBL; S71616; AAB31059.1; -
CC EMBL; U27340; AAB92567.1; -
CC EMBL; U57999; AAB02695.1; -
CC PIR; JH0604; JH0604.
DR MGJ; MGJ:97783; Psap.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapaB_1.
DR InterPro; IPR008138; SapaB_2.
DR InterPro; IPR008139; SaposinB.
DR InterPro; IPR008373; Saposin.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SapaB_1; 4.
DR Pfam; PF03489; SapaB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapaB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPA; 4.
KW Signal; Glycoprotein; Repeat.
FT SIGNAL 1 16
FT CHAIN 17 557
FT DOMAIN 21 54 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 59 142 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 193 277 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 313 394 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 438 519 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 524 557 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 197 273 BY SIMILARITY.
FT DISULFID 200 267 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 317 390 BY SIMILARITY.
FT DISULFID 320 384 BY SIMILARITY.
FT DISULFID 348 359 BY SIMILARITY.
FT DISULFID 442 515 BY SIMILARITY.
FT DISULFID 445 509 BY SIMILARITY.

```

```

RESULT 4
SAP_MOUSE STANDARD; PRT; 557 AA.
AC 061307; 060861; 064219;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN SGP OR SGP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;

```

FT DISULFID 473 484 BY SIMILARITY.
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 83 83 Q -> E (IN REF. 2).
 FT CONFLICT 158 158 I -> V (IN REF. 3).
 FT CONFLICT 160 160 MISSING (IN REF. 2).
 FT CONFLICT 171 172 MS -> SA (IN REF. 3).
 FT CONFLICT 244 244 V -> L (IN REF. 2).
 FT CONFLICT 254 254 M -> I (IN REF. 3).
 FT CONFLICT 255 255 L -> N (IN REF. 3).
 FT CONFLICT 260 260 MISSING (IN REF. 3).
 FT CONFLICT 307 307 F -> D (IN REF. 2).
 FT CONFLICT 322 322 N -> L (IN REF. 2).
 FT CONFLICT 349 350 AL -> GV (IN REF. 1).
 FT CONFLICT 367 367 G -> D (IN REF. 3).
 FT CONFLICT 370 370 L -> Q (IN REF. 2).
 FT CONFLICT 373 373 I -> D (IN REF. 3).
 FT CONFLICT 391 391 A -> T (IN REF. 3).
 FT CONFLICT 393 393 R -> L (IN REF. 3).
 FT CONFLICT 406 406 A -> R (IN REF. 2 AND 3).
 FT CONFLICT 430 430 P -> R (IN REF. 2).
 FT CONFLICT 445 445 C -> F (IN REF. 3).
 FT CONFLICT 448 448 L -> P (IN REF. 4).
 SQ SEQUENCE 557 AA; 61422 MM; 134593E20499E35E CRC64;

Query Match 69.5%; Score 1916.5; DB 1; Length 557;
 Best Local Similarity 63.5%; Pred. No. 2.8e-118;
 Matches 350; Conservative 78; Mismatches 88; Indels 35; Gaps 3;

QY 2 ASLIGALAGPVLGKECTRGASAWCQNVKTAASDCGAYHGLQTVWKNPKYKSPDCDCK 61
 DB 8 ASLIGALAGPVLGKECTRGASAWCQNVKTAASDCGAYHGLQTVWKNPKYKSPDCDCK 67
 QY 62 DVATAGDMKDNATEEELVYLEKTCQDMLPKPMNASCKEIVSYLPIVLIIDIKGEMSR 121
 DB 68 TVTETAGNULKDNAROEELIHLTEKTEWINDSSASCKEVDSDYLVILLDMKGEKSMN 127
 QY 122 PGEVSAINLCESLOKHLAEINHOLESNKIPELDMEVAVAFPMANIPILLYFODGPRS 181
 DB 128 PGEVSAINLCESLOKHLAEINHOLESNKIPELDMEVAVAFPMANIPILLYFODGPRS 186
 QY 182 KPOPDNDVGDCCDQWTDIOTARTNSTVQALVEHVEKECRLGPMADICKYISQ 241
 DB 187 QPQPMADVQDCCKLVSDVQTAVKTNSSFTQGVHDVHCKDCRLGPMADICKYISQ 246
 QY 242 YSEIAIQMMEM---QPKICALVGFCEVEKEMQTLVPAKVASKNVPIVLEVEPIKK 298
 DB 247 YSEVGVQMLMHOQOPKEICVLAGFCNEVAKVPMKTLVPAETITIKNILLPALEMDYEQ 306
 QY 299 HEVPAKSDVGVCEVEELVKEVTKLIDNNKTEKELIDAFDKMCKSLPKSLSEECQVVDY 358
 DB 307 NLVQNHNVITLCQTCQFVWKNKSEELIVNNAVEELLVKGLSNACALLPDPARTKCEVGTG 366
 QY 359 GSSILSILBEVSEPVLCMLTCSG-----TRL 387
 DB 367 GPSLIDFIFHEVNVSSLCGVIGLCARPVELVALEQAPALVSLAKEPPIPPQAPQK 426
 QY 388 PALTIVHTQPKDGFCEVCKLVGLYLDRLNENKSTKOEILAALEKSGSFLPDYQKQCDQ 447
 DB 427 SALPAHVPPQKNGFCVCKLVLYLEHNEKNSTKEELIILAELKSGSFLPDYQKQCDQ 486
 QY 448 FVAEYEPVLEIIVWMDPSFVLKIGAPSAHNPILGTEKCIWGBSYNQMNTETAQCN 507
 DB 487 FVAEYEPVLEIIVWMDPSFVLKIGAPSAHNPILGTEKCIWGBSYNQMNTETAQCN 546
 QY 508 AVECHCKHWN 518
 DB 547 AVDHCKHWN 557

RESULT 5

SAP_CHICK
 ID SAP_CHICK STANDARD, PRT, 518 AA.
 AC 01035;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Proactivator polypeptide precursor [Contains: Saposin A; Saposin B; Saposin C; Saposin D].
 GN PSAP.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 194-203.
 RC TISSUE=Brain, and Liver;
 RX MEDLINE=98129745; PubMed=9461526;
 RA Azuma N., Seo H.-C., Iie O., Pu Q., Gould R.M., Hiraiwa M., Burt D.W.,
 RA Paton I.R., Morrice D.R., O'Brien J.S., Kishimoto Y.,
 RT "Cloning, expression and map assignment of chicken prosaposin.";
 RN Biochem. J. 330:321-327(1998).
 [2]
 RP SEQUENCE FROM N.A.
 RA Altman N., Horowitz M.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: The lysosomal degradation of sphingolipids takes place
 CC by the sequential action of specific hydrolases. Some of these
 CC enzymes require specific low-molecular mass, non-enzymic proteins:
 CC the sphingolipid activator proteins (coproteins) (By similarity).
 CC -1- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
 CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
 CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
 CC Saposin C apparently acts by combining with the enzyme and acidic
 CC lipid to form an activated complex, rather than by solubilizing
 CC the substrate (By similarity).
 CC -1- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
 CC gangliosides by beta-galactosidase A (EC 3.1.6.8), GM1
 CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
 CC Saposin B forms a solubilizing complex with the substrates of the
 CC sphingolipid hydrolases (By similarity).
 CC -1- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
 CC activator (EC 3.1.4.12) (By similarity).
 CC -1- SUBUNIT: Saposin B is a homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: lysosomal (By similarity).
 CC -1- PTM: This precursor is proteolytically processed to 4 small
 CC peptides, which are similar to each other and are sphingolipid
 CC hydrolase activator proteins (By similarity).
 CC -1- SIMILARITY: Contains 2 saposin A-type domains.
 CC -1- SIMILARITY: Contains 4 saposin B-type domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL, AB004471; BAA19914.1; -;
 DR EMBL, AF108656; AAF05899.1; -;
 DR InterPro, IPR003119; SAPA.
 DR InterPro, IPR007856; SAPB.1.
 DR InterPro, IPR008138; SAPB.2.
 DR InterPro, IPR008140; SAPB_sub.
 DR InterPro, IPR008373; Saposin.
 DR InterPro, IPR008139; SaposinB.
 DR Pfam, PF02199; SAPA; 2.
 DR Pfam, PF05184; SAPB.1; 4.
 DR Pfam, PF03489; SAPB.2; 4.
 DR PRINTS, PR01797; SAPOSIN.
 DR ProDom, PD001732; SAPB_sub; 3.

KW Signal, Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat;
 KW GM2-gangliosidosis.
 FT SIGNAL 1 17 POTENTIAL.
 FT PROPER 18 60
 FT CHAIN 61 143 SAPOSIN A.
 FT PROPER 145 193
 FT CHAIN 194 276 SAPOSIN B.
 FT PROPER 278 305
 FT CHAIN 307 387 SAPOSIN C.
 FT PROPER 389 398
 FT CHAIN 399 480 SAPOSIN D.
 FT PROPER 482 518
 FT DOMAIN 22 55 SAPOSIN-LIKE TYPE A 1.
 FT DOMAIN 60 143 SAPOSIN-LIKE TYPE B 1.
 FT DOMAIN 193 277 SAPOSIN-LIKE TYPE B 2.
 FT DOMAIN 307 388 SAPOSIN-LIKE TYPE B 3.
 FT DOMAIN 399 480 SAPOSIN-LIKE TYPE B 4.
 FT DOMAIN 485 518 SAPOSIN-LIKE TYPE A 2.
 FT DISULFID 64 139 BY SIMILARITY.
 FT DISULFID 67 133 BY SIMILARITY.
 FT DISULFID 95 107 BY SIMILARITY.
 FT DISULFID 197 273 BY SIMILARITY.
 FT DISULFID 200 267 BY SIMILARITY.
 FT DISULFID 229 240 BY SIMILARITY.
 FT DISULFID 311 384 BY SIMILARITY.
 FT DISULFID 314 378 BY SIMILARITY.
 FT DISULFID 342 353 BY SIMILARITY.
 FT DISULFID 403 476 BY SIMILARITY.
 FT DISULFID 406 470 BY SIMILARITY.
 FT DISULFID 434 445 BY SIMILARITY.
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 94, 94 R -> T (IN REF. 2).
 FT CONFLICT 486 486 B -> D (IN REF. 2).
 SQ SEQUENCE 518 AA; 57601 MW; B803000E891C3963 CRC64;
 Query Match 60.8%; Score 1676.5; DB 1; Length 518;
 Beot Local Similarity 57.6%; Pred. No. 1.3e-102;
 Matches 300; Conservative 96; Mismatches 112; Indels 13; Gaps 5;

QY 478 SAKRPLCTEKICWGPSYWCNTERRAQCNAVEHCKRHVN 518
 DB 478 AAKRPLGSDACVWGPGYWCXNMETPAQCNAVDHCRHVN 518
 RESULT 6
 SAP_PIG STANDARD; PRT; 80 AA.
 AC P81405;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Saposin B (Cerebroside sulfate activator) (CS-ACT) (Non-specific activator) (Sphingolipid activator protein 1) (SAP-1).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxId=9823;
 RN [1]
 RP SEQUENCE OF 1-79.
 RC TISSUE=Kidney;
 RX MEDLINE=93229506; PubMed=8471613;
 RA Stevens R.L., Fauli K.F., Conklin K.A., Green B.N., Fluharty A.L.;
 RT "Porcine cerebroside sulfate activator: further structural characterization and disulfide identification.";
 RL Biochemistry 32:4051-4059 (1993).
 RN [2]
 RP SEQUENCE OF 1-64.
 RC TISSUE=Kidney;
 RX MEDLINE=9222651; PubMed=1562358;
 RA Fluharty A.L., Katona Z., Meek W.B., Frei K., Fowler A.V.;
 RT "The cerebroside sulfate activator from pig kidney: purification and molecular structure.";
 RL Biochem. Med. Metab. Biol. 47:66-85 (1992).
 RN [3]
 RP STRUCTURE OF CARBOHYDRATE ON ASN-21.
 RX MEDLINE=21110404; PubMed=11180632;
 RA Fauli K.F., Johnson J., Kim M.J., To T., Whitelegge J.P., Stevens R.L., Fluharty C.B., Fluharty A.L.;
 RT "Structure of the asparagine-linked sugar chains of porcine kidney and human urine cerebroside sulfate activator protein.";
 RL J. Mass Spectrom. 35:1416-1424 (2000).
 RN [4]
 RP MASS SPECTROMETRY.
 RC TISSUE=Kidney;
 RX MEDLINE=99441404; PubMed=10510427;
 RA Fauli K.F., Whitelegge J.P., Higginson J., To T., Johnson J., Krutchenky A.N., Standing K.G., Waring A.J., Stevens R.L., Fluharty C.B., Fluharty A.L.;
 RT "Cerebroside sulfate activator protein (Saposin B): chromatographic and electrospray mass spectrometric properties.";
 RL J. Mass Spectrom. 34:1040-1054 (1999).
 CC -I- FUNCTION: Saposin B stimulates the hydrolysis of galacto-cerebrosides by beta-galactosidase A (EC 3.2.1.23) and globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
 CC Saposin B forms a solubilizing complex with the substrates of the sphingolipid hydrolases.
 CC -I- SUBUNIT: Saposin B is a homodimer (by similarity).
 CC -I- PMW: The one residue extended Saposin B-Val is only found in a minority of the chains.
 CC -I- SIMILARITY: Contains 1 saposin B-type domain.
 DR GlycosultedB; P81405; -;
 DR InterPro; IPR007856; SAPB_1.
 DR InterPro; IPR008138; SAPB_2.
 DR InterPro; IPR008373; Saposin.
 DR InterPro; IPR008139; SaposinB.
 DR Pfam; PF05184; SAPB_1; 1.
 DR Pfam; PF03489; SAPB_2; 1.
 DR PRINTS; PR01797; SAPOSIN.
 DR SMART; SM00118; SAPB; 1.
 KW Glycoprotein; sphingolipid metabolism.

FT CHAIN 1 79 SAPOSIN B:
 FT CHAIN 1 80 SAPOSIN B-VAL.
 FT DOMAIN 1 80 SAPOSIN-LIKE TYPE B.
 FT DISULFID 4 77
 FT DISULFID 7 71
 FT DISULFID 36 47
 FT CARBOHYD 21 21
 SQ SEQUENCE 80 AA; 8949 MW; EF7BA29B63E789C CRC64;
 N-LINKED (GLCNAC...) (COMPLEX).
 /FTID-CAR.0001777
 Query Match 14.2%; Score 393; DB 1; Length 80;
 Best Local Similarity 88.8%; Pred. No. 1.4e-19;
 Matches 71; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 QY 189 GDVQCDCIQWVTTIQAVRTNSTFVQALVHVEKECDRLGPGMADICKYISYSRIAIQ 248
 DB 1 GDVQCDCIQWVTTIQAVRTNSTFVQALVHVEKECDRLGPGMADICKYISYSRIAIQ 60
 QY 249 MAMHMQPKETICALVGFCEV 268
 DB 61 MAMHMQPKETICALVGFCEV 80
 RESULT 7
 PSPB_MOUSE STANDARD; PRT; 377 AA.
 AC P50405;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pulmonary surfactant-associated protein B precursor (SP-B) (Pulmonary
 DE surfactant-associated proteolipid SPL(Phc)).
 GN SPTPB OR SPTP3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=BA/2J; TISSUE=Liver;
 RX MEDLINE=95208782; PubMed=7900819;
 RA Bruno M.A., Bohinski R.J., Carter J.E., Fosse K.A., Wiltset J.A.;
 RT "Structure and function of the mouse surfactant protein B gene.";
 RL Am. J. Physiol. 268:1381-1389(1995).
 CC -1- FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE
 CC LIQUID INTERFACE BY LOWERING THE SURFACE TENSION AT THE AIR-
 CC THE COLLAPESE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS
 CC PER METER.
 CC -1- SUBUNIT: Homodimer; disulfide-linked.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipids and 10%
 CC protein. There are 4 surfactant-associated proteins: 2 collagenous,
 CC hydrophobic-binding glycoproteins (SP-A and SP-D) and 2 small
 CC -1- SIMILARITY: Contains 1 saposin A-type domain.
 CC -1- SIMILARITY: Contains 3 saposin B-type domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 DR EMBL: S78114; AAB34846.2; -;
 DR HSSP: P07988; 1DFW.
 DR MGD; MG1:109516; Sftpb.
 DR InterPro: IPR003119; Sapa.
 DR InterPro: IPR007856; Sapa.1.
 DR InterPro: IPR008138; Sapa.2.
 DR InterPro: IPR008140; Sapa_Sub.

DR InterPro: IPR008373; Saposin.
 DR InterPro: IPR008139; SaposinB.
 DR InterPro: IPR008137; Surfactant_B.
 DR Pfam: PF02199; SABA.1.
 DR Pfam: PF05184; SABA_1; 1.
 DR Pfam: PF03489; SABA_2; 3.
 DR PRINTS: PR01797; SAPOSIN.
 DR ProDom: PD001732; Sapa_Sub; 1.
 DR ProDom: PD008002; Surfactant_B; 1.
 DR SMART: SM00162; SABA.1.
 DR SMART: SM0018; SABA; 3.
 DR Surface Film; Gaseous exchange; Glycoprotein; Repeat.
 FT CHAIN 1 191
 FT PROPEP 271 377 PULMONARY SURFACTANT-ASSOCIATED PROTEIN
 FT DOMAIN 27 60 B.
 FT DISULFID 199 268 SAPOSIN-LIKE TYPE A.
 FT DISULFID 202 262 BY SIMILARITY.
 FT DISULFID 226 237 BY SIMILARITY.
 FT DISULFID 239 239 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 307 307 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 377 AA; 41728 MW; CB687A82BA3FC56C CRC64;
 Query Match 13.7%; Score 378; DB 1; Length 377;
 Best Local Similarity 22.5%; Pred. No. 9.5e-18;
 Matches 111; Conservative 73; Mismatches 145; Indels 164; Gaps 16;
 QY 18 ECTGSAVWQCVKNTSDGAVVHCGIQVWNNKPTVSLPCDICKVYTAAGMADKNATE 77
 DB 30 ECAQGPQPCQSGSEHVAQCRAHCHQAEVWGHAGANDL-CQCEIDVHLITTKTEDEAQ 88
 QY 78 EELVTEKTCQMLPFPNNSASCKEIVSYLVIIIDIKGMSRPGVCSALNLCESLQK 137
 DB 89 EAIRKFLBECEDILPKLIVPRQGVLDVYLPLVIDYFQSQIN-PKALCNHGLC----- 142
 QY 138 HIAELNHQQLSBNKIPELDMTEVAVPFPANIPLLIYPODGRSKPQPKDNGVQDCIQ 197
 DB 143 -----PRGAKPEON----- 152
 QY 198 MYTDIQAVRTNSTFVQALVHVEKECDRLGPGMADICKYISYSRIAIQWMMHMQPK 257
 DB 153 -----PGMD----- 157
 QY 258 ICALVGFCEVEMEMQTLVPKAKSKVTPALE-LVEP-----IKKHEVPAKSDVY 308
 DB 158 -----AVPNPLLDKLVLPVLPALALARGPHTQDPSEQQLPIPLP-F 198
 QY 309 CEVCEFLVETKTLIDNNKTEKEILD-AFDKMKSLPKSLSECOEVDVDTYGGSLITLL 367
 DB 199 CWLCRTLLIKRYQAVI-----PKGVLAIVASQVCHVPLVVGICQCLAERTVLLDL 253
 QY 368 EYSEVELYCSMLHUCS-----GTRLPALTVAHQ-PKDGFCVCEVKLVGLIDNLEKNS 421
 DB 254 GRVIVQLVCGVILRCSTEDAMGPALEVEPIIEPPIQDTCHRCVKSI-----NQAMNT 308
 QY 422 TKQETIALLEKGC-SFLDDPYQKQCDQVFAEYEVLEILVENVDPFVCLIKTAGCSAH 480
 DB 309 SEQAMPQAMHQAQLRFMTD-RQKCEQFVEQHMQLLALVRSODAHITCOALGVCEAPA 366
 QY 481 KPLGTEKCIWGP 493
 DB 367 SPL-----QCFQTP 375
 RESULT 8
 PSPB_RAT STANDARD; PRT; 376 AA.
 AC P22355;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pulmonary surfactant-associated protein B precursor (SP-B) (Pulmonary

DE surfactant-associated proteolipid SPL(Phe)).

GN SFRPB OR SFTP3.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]

RN SEQUENCE FROM N.A.

RP MEDLINE=89150284; PubMed=2920185;

RA Emrie P.A., Shannon J.M., Mason R.J., Fisher J.H.,
"cDNA and deduced amino acid sequence for the rat hydrophobic
pulmonary surfactant-associated protein, SP-B,"
Biochim. Biophys. Acta 994:215-221(1989).

CC -! FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE
ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-
LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES
THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS
PER METER.

CC -! SUBUNIT: Homodimer; disulfide-linked.

CC -! SUBCELLULAR LOCATION: Extracellular.

CC -! MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
protein. There are 4 surfactant-associated proteins: 2 collagenous,
carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small
hydrophobic proteins (SP-B and SP-C).

CC -! SIMILARITY: Contains 1 saposin A-type domain.

CC -! SIMILARITY: Contains 3 saposin B-type domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X14778; CAA32885.1; -

DR PIR; S02766; S02766.

DR HSSP; P07988; 1DFW.

DR InterPro; IPR003119; Sapa_1.

DR InterPro; IPR007856; Sapa_1.

DR InterPro; IPR008138; Sapa_2.

DR InterPro; IPR008140; Sapa_sub.

DR InterPro; IPR008373; Saposin.

DR InterPro; IPR008139; SaposinB.

DR InterPro; IPR008137; Surfactant_B.

DR Pfam; PF02119; Sapa_1.

DR Pfam; PF05184; Sapa_1; 1.

DR Pfam; PF03489; Sapa_2; 3.

DR PRINTS; PR01797; SAPOSIN.

DR ProDom; PD001732; Sapa_sub; 1.

DR ProDom; PD008002; Surfactant_B; 1.

DR SMART; SM00162; SAPA; 1.

DR SMART; SM00118; SAPB; 3.

DR Surface film; Gaseous exchange; Glycoprotein; Repeat.

KW Surface film; Gaseous exchange; Glycoprotein; Repeat.

FT PROPEP 1 190

FT CHAIN 191 269

FT PULMONARY SURFACTANT-ASSOCIATED PROTEIN

FT B.

FT PROPEP 270 376

FT DOMAIN 26 59

FT DISULFID 198 267

FT DISULFID 201 261

FT DISULFID 225 236

FT DISULFID 228 238

FT CAROHD 306 306

FT SEQUENCE 376 AA; 41590 MW; F329DC62E733FB4C CRC64;

Query Match 13.5%; Score 372; DB 1; Length 376;
Best Local Similarity 22.7%; Pred. No. 2.3e-17;
Matches 115; Conservative 74; Mismatches 151; Indels 166; Gaps 17;

QY 5 LGAALAGVLTGLKECTRGSAVWCQVWKTASDCGAVHCLQTVWNNPVTYSLPDCIDCKDV 64

DB 18 LGAAATES--ASSPDCAGPKFWCQSLQELICRALGHCLQEVWGAANDL-CQCECDIV 74

QY 65 TAAGDMLKDNATEBEILLVEKTCQDMLPKPMMSASCKEIVDSYLPVILDIKGEWRPGE 124

DB 75 HLRTMTKEDAPFQDITRKFRQECQDILPLKLVPRCROVLYPLVDFYQGGI-KRKA 133

QY 125 VCSALNLCESLQKHLAEINHQKQLESNKIPELDMTVEVAPPMANIPLLLYPDGGRSRQ 184

DB 134 ICSHVGLC-----PL---GQTRPQKPE 153

QY 185 PKNDGVQDCQICQMTVDIQTAVRTNSFTVQALVEHVKECDRLGKADICKNTISQYSE 244

DB 154 -----

QY 245 IAIQMMHMQKEICALVFCDEVKEMQGLVPAKVAKNVIPALE---LVEP----- 295

DB 154 -----

QY 296 IKHSEVPAKSDVYCEVECEFLVKEVYKLIIDNNKTEKILID-AFDKMSKTLPKSLSECOEV 354

DB 186 LSEQQLPIPL-FCMLCRTLLIKRQAVI-----PKGVLAIVSQCHVPLVVGICQCL 239

QY 355 VDTGSSILSLILBEVSEPLVCSMLHCS-----GTRLPALTHTVQ-PKDGFCVECKK 408

DB 240 AERYTVLLDLALGRVVPQVLCGLVRCSTADAIQPALPALEPILEKWPLODTECHFKCS 299

QY 409 LVGYLDRLKNSRKQETILALEKGC-SFLDPYQKODQDVAEVEPTLILIVEVMPDS 467

DB 300 VI-----NQAWNTSQAMPQAMHQACRFWLD--RQKEQFVEQHPQLLALVPRSDAH 352

QY 468 FVCLKIGACPSAHPKPLLTGTEKCIWCP 493

DB 353 TSCQALGVCEAPASPL-----QCFQTP 374

RESULT 9

PSPB HUMAN

ID PSPB HUMAN STANDARD; PRT; 381 AA.

AC P07988; Q96R04;

DT 01-AUG-1988 (Rel. 08, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Pulmonary surfactant-associated protein B precursor (SP-B) (6 kDa
protein) (Pulmonary surfactant-associated proteolipid SPL(Phe)) (18
kDa pulmonary surfactant protein).

GN SFRPB OR SFTP3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89170128; PubMed=2924687;

RA Pilot-Matias T.J., Kister S.B., Fox J.L., Kropp K., Glaeser S.W.,
Whitsett J.A.;

RT "Structure and organization of the gene encoding human pulmonary
surfactant proteolipid SP-B,"
J. Biol. Chem. 262:9808-9811(1987).

RL DNA 8:75-86(1989).

RL (2)

RN SEQUENCE FROM N.A., AND SEQUENCE OF 201-214.

RP TISSUE=Lung;

RX MEDLINE=87250653; PubMed=3597440;

RA Jacobs K.A., Phelps D.S., Steindink R., Fisch J., Kriz R.,
Mitscock L., Dougherty J.P., Taensch H.W., Floros J.;

RT "Isolation of a cDNA clone encoding a high molecular weight precursor
to a 6-kDa pulmonary surfactant-associated protein,"
J. Biol. Chem. 262:9808-9811(1987).

RL [3]

RN SEQUENCE FROM N.A., AND VARIANTS ILE-131; PHE-176 AND HIS-272.

RP Rieder M.J., Carlington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
Nickerson D.A.;

RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

	Query Match	12.3%	Score 340.5:	DB 1:	Length 381;		
	Beset Local Similarity	21.4%;	Pred.	No.2.7e-15;			
	Matches	109;	Conservative	71;	Mismatches	169;	
				Indels	161;	Gaps	16
Oy	1	LASLLGALAGAVLDELKECTRGSAAWCQVWVKTPADCGAAYKICIQITWYNNKPYSLEPCDIC	60				
Db	14	LPTLCGPETAAMTTSLSLACAQCFEWCSCSLLEGALGCRAIGHCLGLQEVWGIVAGDDL-CQCIC	72				
Oy	61	KDVVTAAQDMLKDNATTEELIYLKTCIDMLPKPNMSASCKEIIVDSYLEFVLIIIGKENS	120				
Db	73	EDIVHILNKMAEKAEIFODTRMKFEQECSNVLPLKLIMPQCNOGVLDYDFPVLVIDYFQNGYTD	132				
Oy	121	RPGVCSTANI.CESIOXKHAEINHQKOLESNKIPELDMTEVVAFPFANIPLLLYPDGGR	180				
Db	133	SNG--ICMHLGIICKSRQ-----	147				
Oy	181	SKPOPKNDGVQCDCIQMTVDIQTAFTNSTFVQALVEHKECDEBLGPMADICKNYIS	240				
Db	148	--PEREQ-----	158				
Oy	241	QYSEIAIQMMHMOPKEICALVFCEDEVKEMQTLVPAPKVASKNVIIPALEIVEPIKHE	300				
Db	159	-----PLRK-----	189				
Oy	301	VPAKSDV-----YCEVCEFLVKEYTKLIDNNKTBEKELIDAFDKKCSLKPSLSSE	350				
Db	190	GPHOTDLSEQQPFIPLPYCMWLCPRALIKIRIQAMIPIKG---ALAVAVALCVRPVPLVAGSI	245				
Oy	351	COEVDVTFGSSLSILILEVSPELVCSMI.HCS-----CTR.PALTVHTQPKDGSGFCBV	405				
Db	246	CQCLAERSVILIIDLTLGRMLPOLYCRLVLRCSMDSDSAPRSP---TGSM.LPRDSE-CHL	301				
Oy	406	CKKLVGYIDRLNLKNSTKQGEILAALEKGC--SF.LDPDYOKOCDOFAVABEPVLIBELIVEV	463				
Db	302	CMSVY-----TQAGNSSBQAIPQMALQACVGSWLD---REKKQGFQEGHTPQLLTIVRPG	353				
Oy	464	MDPSFVCLKIGACPSAHKPLLTEKCINCIP	493				
Db	354	WDHHTTCCALGIVCGTMSSPL-----QCIHSP	379				
<hr/>							
	RESULT 10						
	PSPB CANFA	STANDARD;	PRT;	363	AA.		
AC	P17129;						
DT	01-AUG-1990 (Rel. 15, Created)						
DT	01-AUG-1990 (Rel. 15, Last sequence update)						
DT	16-OCT-2001 (Rel. 40, Last annotation update)						
DE	Pulmonary surfactant-associated protein B precursor (SP-B) (6 kDa protein) (pulmonary surfactant-associated proteolipid SPL(Phe))						
DE	(Pulmonary surfactant protein 18) (SP 18) (Fragment).						
GN	SFPB OR SFTPB.						
OS	Canis familiaris (dog).						
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
CC	Mammalia; Eutheria; Carnivora; Fissipedida; Canidae; Canis.						
OX	NCBI_TaxId=9615;						
RN	[1]						
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 182-211.						
RC	TISSUE=Lung;						
FX	MEDLINE=87092398; PubMed=3467361;						
RA	Havgood S., Benson B.J., Schilling J., Damm D., Clements J.A.,						
RT	White R.T.;						
RT	"Nucleotide and amino acid sequences of pulmonary surfactant protein						
RT	SP 18 and evidence for cooperation between SP 18 and SP 28-36 in						
RL	surfactant lipid adsorption.";						
CC	Proc. Natl. Acad. Sci. U.S.A. 84:66-70(1987).						
CC	- FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE						
CC	ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-						
CC	LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES						
CC	THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS						
CC	PER METER.						
CC	- SUBUNIT: Homodimer; disulfide-linked.						
CC	- SUBCELLULAR LOCATION: Extracellular.						

```

CC -1- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
CC protein. There are 4 surfactant-associated proteins: 2 collagenous,
CC carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small
CC hydrophobic proteins (SP-B and SP-C).
CC -1- SIMILARITY: Contains 1 saposin A-type domain.
CC -1- SIMILARITY: Contains 3 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M15170; AAA30893.1; -.
DR PIR; B29072; A29072.
DR HSSP; P07988; IDEW.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; Sapa_1.
DR InterPro; IPR008138; Sapa_2.
DR InterPro; IPR008140; Sapa_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR InterPro; IPR008137; Surfactant_B.
DR Pfam; PF02199; SAPA; 1.
DR Pfam; PF05184; Sapa_1; 1.
DR Pfam; PF03489; Sapa_2; 3.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; Sapa_sub; 1.
DR ProDom; PD008002; Surfactant_B; 1.
DR SMART; SM00162; SAPA; 1.
DR SMART; SM00118; SAPA; 3.
KW Surface film; Gaseous exchange; Glycoprotein; Repeat.
FT PROPEP 1 1
FT FT <1 180
FT FT 181 259
FT CHAIN
FT
FT PROPEP 260 363
FT DOMAIN 21 54
FT DISULFID 188 257
FT DISULFID 191 251
FT DISULFID 215 226
FT DISULFID 228 228
FT CARBOHYD 293 293
SQ SEQUENCE 363 AA; 40180 MW; P4DAD0E02D982719 CRC64;

Query Match 11.1%; Score 305; DB 1; Length 363;
Best Local Similarity 21.6%; Pred. No. 5,4e-13;
Matches 108; Conservative 75; Mismatches 159; Indels 158; Gaps 19;

OY 5 LGAA-LAGPVLGLKECTRGSAVWQONKTKASDGAIVHGIQOTVWNNKRTYVSLPEDICKDV 63
Db 13 LGAADWAPSLSA--CARGPRFWQCSLEQALQCCALGHCIQEWGNARADDL-CQECODI 68
OY 64 VTAAGDMILKDNATEEELIVYLEKTCDMILPKPNNSASCKEIVDSYLPYLLDIIKEMSRPG 123
Db 69 VRILTNTKTKAIPQDMWRKRFLEHNCVDLPRLKTLTPQCHNHMGITFPVVDVFFQGIN-PK 127
OY 124 EVCSALNLCSLQKHIALNLHQQKQLESNKIPELDMTEVAVAPFMANIPDLLLYPDGPRSKP 183
Db 128 IICGHLGLC--KPGLEPERQESLSDPLDLKILPEL----- 162
OY 184 QPKNGEIVQDCDIOMTVDTIDQPARVNTSTFQALVAENHKEBCDRLGPGMADICKNYISOYS 243
Db 163 -----PALQVLTGPHTDLS 177
OY 244 EIALQMMHMQPKKICALVGFCDCEKMPQOTLPAKAVASKNVLPALVELPIKQHEVPA 303
Db 178 E-----QQLP-----LP- 186
OY 304 KSDVYCEVCEFLVKEVTKLIDNNKTEKILD-AFDKNCCKLPSLSSECCQEVVDYGGSI 362

```

```

Db 187 -----YCWLCRTLIKRIQAMI-----PKGVLAIVTGVGVHVVPLVVGICQCLGRYVLL 237
Cc 363 LSILLEEVSPELVCSMLHCS-----GTRLPALVTHTVTPRODGFCEVCKKLVGYLDRNL 417
Cc 238 LDALLGRLPQVGVGLVRCSHEDSAGPALSLPSE-WSRQESK-CQLCMFVT-----TQ 230
Cc 418 EKNSKTQKEIILALEKGC--SFLPDPYQKQCDQFAVEYEPVLIETLVWMPSPVCLKIGA 475
Cc 291 AGNSHQETPQAIHQACILSSWLD---RQKCEQFVECHMPRLQTLASGRDHTTCQALGA 347
Cc 476 CPANHKPLLTGTEKIMKPSY 495
Cc 348 CRTTFSP-----QCHITHP 363

RESULT 11
PSPB RABIT
ID PSPB RABIT STANDARD; PRT; 370 AA.
AC P15285;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pulmonary surfactant-associated protein B precursor (SP-B) (6 kDa protein) (Pulmonary surfactant-associated proteolipid SPL(phe)).
GN SFTPB OR SFTPB3
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=89228033; PubMed=2469419;
RA Xu J., Richardson C., Ford C., Spencer T., Li-Juan Y., Mackie G.,
RT Hammond G., Possmayer F.;
RT "Isolation and characterization of the cDNA for pulmonary surfactant-associated protein-B (SP-B) in the rabbit."
RL Biochem. Biophys. Res. Commun. 160:325-332(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white;
RX MEDLINE=95208794; PubMed=7900830;
RA Maxgana R.K., Boggaram V.;
RT "Transcription and mRNA stability regulate developmental and hormonal expression of rabbit surfactant protein B gene."
RL Am. J. Physiol. 268:1481-1490(1995).
RN [3]
RP SEQUENCE OF 1-34 FROM N.A.
RX MEDLINE=96096536; PubMed=8522191;
RA Luzi P., Anceleschi M., Strayer D.S.;
RT "The upstream region of the SP-B gene: intrinsic promoter activity and glucocorticoid responsiveness related to a new DNA-binding protein."
RL Gene 165:285-290(1995).
Cc -1- FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS PER METER.
Cc -1- SUBUNIT: Homodimer; disulfide-linked.
Cc -1- SUBCELLULAR LOCATION: Extracellular.
Cc -1- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10% protein. There are 4 surfactant-associated proteins: 2 collagenous, carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small hydrophobic proteins (SP-B and SP-C).
Cc -1- SIMILARITY: Contains 1 saposin A-type domain.
Cc -1- SIMILARITY: Contains 3 saposin B-type domains.
Cc -----
Cc This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

```

```

Cc entitles requires a license agreement (See http://www.eib.ch/announce/
Cc or send an email to license@ib.eib.ch).
Cc -----
Cc EMBL, M24901; AAA31466.1; -
Cc EMBL, U17106; AA67934.1; -
Cc EMBL, S80649; AADI435.1; -
Cc PIR, A32421; LNRB.
Cc PIR, I46531; I46531.
Cc HSSP, P07988; IDPM.
Cc InterPro; IPR003119; Sapa.
Cc InterPro; IPR007856; Sapa_1.
Cc InterPro; IPR008138; Sapa_2.
Cc InterPro; IPR008140; Sapa_sub.
Cc InterPro; IPR008137; SaposinB.
Cc Pfam; PF02199; SABA; 1.
Cc Pfam; PF05184; Sapa_1; 1.
Cc Pfam; PF03489; Sapa_2; 3.
Cc ProDom; PD001732; Sapa_sub; 1.
Cc ProDom; PD008002; Surfactant_B; 1.
Cc SMART; SM00162; SABA; 1.
Cc SMART; SM00118; SABA; 3.
Cc Surface film; Gaseous exchange; Glycoprotein; Repeat.
Cc PROPEP
Cc CHAIN
Cc 1 184
Cc 185 263
Cc 264 370
Cc 29 62
Cc 192 261
Cc 195 255
Cc 219 230
Cc 232 232
Cc 300 300
Cc 15 15
Cc 184 184
Cc 232 232
Cc 329 355
Cc 40610 MW; 42304769512DCBS CRC64;
Cc VQGLGCRNLRPEPGRV (IN REF. 1).
Cc -----
Cc Query Match 10.5%; Score 290; DB 1; Length 370;
Cc Best Local Similarity 21.6%; Pred. No. 5.3e-12;
Cc Matches 111; Conservative 65; Mismatches 167; Indels 170; Gaps 19;
Cc -----
Cc 1 LSLIGALAGAVLDL-----KECTGSAVWCQVKTASDCAVYHCLQTVWKPRTYKSL 55
Cc 10 LLLLLPFLCGFTAVWATSPPLACQGEFECOSLEQALCKALGHCIQEVWGHVADDL 69
Cc 56 PCDDICKDVVTAAGMLKDNATEEELIVLEKTCMLPKPNNASASCKEIVSYLPVLIIT 115
Cc 70 -CQECQDIVNLTWKTKKAIQODITRKLEHCECVLPKILVPCQHHVLDVYFPLTTFY 128
Cc 116 KQEMRPREVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAAPMANIPLLLYP 175
Cc 129 OSQINAKA-ICQHGGLCO-----PGSEPEPLD-----PLPKVLVPTLL- 166
Cc 176 QDGPESKQPPQNDGVQDCIQVWTDIQTVRNTSTVQVAVHVEKSCRLRPGNADIC 235
Cc 167 -GALPAK-----SPHTDLS 181
Cc 236 KNYISQISEAIQMMEMQPKICALVGFDEVEMQMOTLVPAKVASKNVLPALIEVER 295
Cc 182 -----AQRPIPL-----P 190
Cc 296 IKKEHVAKSDVYGEVEFLVKEVTKLIDNNKTEKEILD-APDKMCKSLPKSLSECOEV 354
Cc 191 L-----CWLCTLIKRIQAMI-----PKGVLAIVTGVGVHVVPLVVGICQCL 233
Cc 355 VDTYGSILSILLEEVSPELVCSMLHCS-----GTRLPALVTHTVTPRODGFCEVCK 407
Cc 234 AERTYVILLVLLGVHLPQVGVGLVRCSSVDSIGVPTLELPEGEMLPDDE-CRLDM 292
Cc 408 KLVGYLDRNLERKSTQKEIILAA-----LEKGSFLPDPYQKQCDQFAVEYEPVLIETLIVE 462

```

Db 293 SVTQQA-RNISEQRPQAVYHACUSQDQK-----GECEQFVELHPTQLSLISR 341
QY 463 VMDPSFVLKIGACSAKPLIGTEKCIWGSY 495
Db 342 GMDARICQALGACVATLSPF-----QCIQSPHF 370

RESULT 12
SAP_CAVPO STANDARD; PRT; 81 AA.
ID SAP_CAVPO STANDARD; PRT; 81 AA.
AC P20097;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Saposin C (CO-beta-glucosidase) (glucosylceramidase activator)
DS (Sphingolipid activator protein 2) (SAP-2).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=89066787; PubMed=3198642;
RA Sano A, Radin N.S., Johnson L.L., Tarr G.E.;
RT guinea pig liver. Improved isolation method and complete amino acid
RT sequence.";
RL J. Biol. Chem. 263:19597-19601(1988).
CC -1- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
CC Saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate.
CC -1- SIMILARITY: Contains 1 saposin B-type domain.
DR PIR: A32026; A32026.
DR HSSP; P42210; 1QDM.
DR InterPro; IPR007856; SApB_1.
DR InterPro; IPR008138; SApB_2.
DR InterPro; IPR008140; SApB_sub.
DR InterPro; IPR008173; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF05184; SApB_1; 1.
DR Pfam; PF03489; SApB_2; 1.
DR PRINTS; PR01797; SApOSIN.
DR PRODOM; PD001732; SApB_sub; 1.
DR SMART; SM00118; SApB; 1.
DR KEGG; S000118; SApB; 1.
FT DOMAIN 1 81 SAPOIN-LIKE TYPE B.
FT CARBOHD 22 22 N-LINKED (GLCNAC. . .).
FT DISULFD 5 78 BY SIMILARITY.
FT DISULFD 8 72 BY SIMILARITY.
FT DISULFD 36 47 BY SIMILARITY.
SQ SEQUENCE 81 AA; 8852 MW; E564CE1F0A292556 CRC64;

Query Match 9.5%; Score 261; DB 1; Length 81;
Best Local Similarity 61.8%; Pred. No. 6.4e-11;
Matches 47; Conservative 18; Mismatches 11; Indels 0; Gaps 0;

QY 307 VYCEVCELVKVKTLIDNNTEKILDAFKMKSKLPKSLSECEQVVDYTGSSILSL 366
Db 3 VTCACACEVAVKQVELLIDNNTEKTIHALDSVCLLPESVSEVCEQVVDYTGSSIVALL 62

QY 367 LEEVSPELVCSMLHLC 382
Db 63 LQEMSPELVCSMLHLC 78

RESULT 13
ASPR_CUCPE STANDARD; PRT; 513 AA.
ID ASPR_CUCPE STANDARD; PRT; 513 AA.

AC 004057;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Aspartic proteinase precursor (EC 3.4.23.-).
OS Cucurbita pepo (Vegetable marrow) (Summer squash).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eudotsids I; Cucurbitales; Cucurbitaceae; Cucurbita.
OX NCBI_TaxID=3663;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97354181; PubMed=9210475;
RA Hiraiwa N, Kondo M, Nishimura M, Hara-Nishimura I;
RT "An aspartic endopeptidase is involved in the breakdown of
RT propeptides of storage proteins in protein-storage vacuoles of
RT plants.";
RL Eur. J. Biochem. 246:133-141(1997).
CC -1- FUNCTION: Involved in the breakdown of propeptides of storage
CC proteins in protein-storage vacuoles.
CC -1- SUBUNIT: Heterodimer of a 32 kDa subunit and a 16 kDa subunit.
CC -1- SUBCELLULAR LOCATION: Vacuolar.
CC -1- SIMILARITY: Belongs to peptidase family A1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB002695; BAA19607.1; -.
DR PIR; T09739; T09739.
DR HSSP; P42210; 1QDM.
DR MEROPS; A01.020; -.
DR InterPro; IPR001969; Asprotease AS.
DR InterPro; IPR009007; Pept_A acid.
DR InterPro; IPR001461; Peptidase_A1.
DR InterPro; IPR007856; SApB_1.
DR InterPro; IPR008138; SApB_2.
DR InterPro; IPR008140; SApB_sub.
DR InterPro; IPR008173; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF00026; asp; 1.
DR Pfam; PF05184; SApB_1; 1.
DR Pfam; PF03489; SApB_2; 1.
DR PRINTS; PR00792; PEBSIN.
DR PRINTS; PR01797; SAPOIN.
DR PRODOM; PD001732; SApB_sub; 1.
DR SMART; SM00118; SApB; 2.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
DR KEGG; S000141; ASP_PROTEASE; 2.
FT SIGNAL 1 25 POTENTIAL.
FT PROPEP 26 73 POTENTIAL.
FT CHAIN 26 513 ASPARTIC PROTEINASE.
FT DOMAIN 322 420 SPECIFIC TO PLANT ASPARTIC PROTEINASES.
FT ACT_SITE 108 108 BY SIMILARITY.
FT ACT_SITE 294 294 BY SIMILARITY.
FT DISULFD 120 126 BY SIMILARITY.
FT DISULFD 285 289 BY SIMILARITY.
FT CARBOHD 432 469 BY SIMILARITY.
FT CARBOHD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 513 AA; 55855 MW; 083FB7064CE02DC2 CRC64;

Query Match 6.3%; Score 173.5; DB 1; Length 513;
Best Local Similarity 29.0%; Pred. No. 0.00034;
Matches 47; Conservative 31; Mismatches 59; Indels 25; Gaps 7;

QY 347 LSECEQVVDYTGSSILSLLEVSPELVCSMLHLC--GTRLPALTVH-VTQPKDG--- 400
Db 320 VSQCKAVVAVQYGGTINDLLSEADPKKICGQINLCTFDGRTGVSWMGISVVDENAGKSS 379

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:25:43 ; Search time 51.6524 Seconds
(without alignments)
3164.197 Million cell updates/sec

Title: US-09-743-684a-1_COPY_7_524
Perfect score: 2759
Sequence: 1 LASLIGALAGPVGLKECT.....NTETAQCNVAVHCRRHYWN 518

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: sp archaea:*
2: sp bacteria:*
3: sp fungi:*
4: sp human:*
5: sp invertebrate:*
6: sp mammal:*
7: sp mhc:*
8: sp organelle:*
9: sp phage:*
10: sp plant:*
11: sp rodent:*
12: sp virus:*
13: sp vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1928	69.9	554	11	Q8BFQ1	Q8BFQ1 mus musculu
2	1437	52.1	512	13	Q7SY70	Q7SY70 xenopus lae
3	1412	51.2	520	13	Q8UVZ4	Q8UVZ4 brachydand
4	1341	48.6	522	13	Q9DGB2	Q9DGB2 brachydand
5	1089.5	39.5	525	11	Q8C1C1	Q8C1C1 mus musculu
6	776	28.1	449	11	Q8BJV5	Q8BJV5 mus musculu
7	772	28.0	402	11	Q8C1N0	Q8C1N0 mus musculu
8	549.5	19.9	953	5	Q9Y125	Q9Y125 drosophila
9	545	19.8	241	4	Q8N7T4	Q8N7T4 homo sapien
10	517	18.7	121	6	P79254	P79254 ovis aries
11	511	18.5	965	5	O15997	O15997 bombyx mori
12	432.5	15.7	876	5	Q8IMH4	Q8IMH4 drosophila
13	343.5	12.5	485	5	Q95X02	Q95X02 naegleria f
14	343.5	12.5	484	5	Q9BMM1	Q9BMM1 naegleria f
15	342.5	12.4	378	11	O35489	O35489 cavia porce
16	330.5	12.0	374	6	Q9TUB1	Q9TUB1 ovis aries

17	327.5	11.9	441 5 Q9UPP3	Q9UPP3 drosophila
18	288	10.4	307 5 Q9BKM2	Q9BKM2 naegleria f
19	286	10.4	370 6 P79333	P79333 oryctolagus
20	279	10.1	294 5 Q95X03	Q95X03 naegleria f
21	279	10.1	456 5 Q94472	Q94472 dictyosteli
22	251	9.1	316 5 Q86PA4	Q86PA4 drosophila
23	220.5	8.0	429 5 Q18276	Q18276 caenorhabdi
24	213	7.7	213 10 Q9SC75	Q9SC75 arabidopsis
25	203.5	7.4	217 10 Q9LZW6	Q9LZW6 arabidopsis
26	187.5	6.8	243 6 Q9T705	Q9T705 ovis aries
27	180	6.5	200 5 Q86JDB	Q86JDB dictyosteli
28	179.5	6.5	240 10 Q9AS89	Q9AS89 oryza sativ
29	176.5	6.4	402 5 Q18279	Q18279 caenorhabdi
30	176	6.4	228 6 Q9BMD9	Q9BMD9 ovis aries
31	173.5	6.3	507 10 Q9ME14	Q9ME14 vitis ripar
32	173.5	6.3	507 10 Q9FRW7	Q9FRW7 nepenthes a
33	172	6.2	188 5 Q86KA8	Q86KA8 dictyosteli
34	171.5	6.2	486 10 Q38934	Q38934 arabidopsis
35	171.5	6.2	506 10 Q65390	Q65390 arabidopsis
36	165.5	6.0	506 10 Q39311	Q39311 brassica na
37	163.5	5.9	273 10 Q9LUX5	Q9LUX5 pyrus pyrif
38	163.5	5.9	513 10 Q8VYJ3	Q8VYJ3 arabidopsis
39	162.5	5.9	514 10 Q41713	Q41713 vigna ungu
40	162	5.9	514 10 Q8L6A9	Q8L6A9 theobroma c
41	161	5.8	107 10 Q8H0R9	Q8H0R9 cynara card
42	157	5.7	512 10 Q04593	Q04593 arabidopsis
43	156.5	5.7	514 10 Q941A2	Q941A2 glycine max
44	155.5	5.6	205 6 Q9N275	Q9N275 ovis aries
45	155.5	5.6	292 10 Q43407	Q43407 brassica ol

ALIGNMENTS

RESULT 1
Q8BFQ1 PRELIMINARY; PRT; 554 AA.
ID Q8BFQ1
AC Q8BFQ1
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Prosaposin.
GN PSAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOB; TISSUE=Kidney, and Thymus;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK086369; BAC40308.1; -;
DR EMBL; AK089998; BAC41035.1; -;
DR MGI; MGI:97783; Psap.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006655; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; SAPA.
DR InterPro; IPR007856; SAPB_1.
DR InterPro; IPR008138; SAPB_2.
DR InterPro; IPR008140; SAPB_sub.
DR InterPro; IPR008173; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SAPB_1; 4.
DR Pfam; PF03489; SAPB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SAPB_sub; 3.
DR SMART; SM00162; SAPA; 2.

DR SMART; SMO0118; SAMP; 4.
SQ SEQUENCE 554 AA; 61050 MW; FF58DB79C7CC0C18 CRC64;
Query March 69.9%; Score 1928; DB 11; Length 554;
Best Local Similarity 63.9%; Pred. No. 3e-138;
Matches 350; Conservative 78; Mismatches 88; Indels 32; Gaps 2;
QY 2 ASLIGALAGPVILGKECTRGSAVWCONVKTASDCGAVHCLQTWNKPRVKSLEPDICK 61
DB 8 ASLILATLSPVQDPKTCSSGSALVCRDVTATVDCGAVHCLQTWNKPRVKSLEPDICK 67
QY 62 DVYTAAGDMLKDNATEBEILIVLEKTCMLPKPNMSASCKEIVDSYLYPIILIKGEMSR 121
DB 68 TVVTEAGNMLKDNATEBEILIVLEKTCMLPKPNMSASCKEIVDSYLYPIILIKGEMSR 127
QY 122 PGEVCSALNLCESLOKHAELNHOKELESNKIPELDMTEVAVPAPNATLILYPODGERS 181
DB 128 PGEVCSALNLCESLOKHAELNHOKELESNKIPELDMTEVAVPAPNATLILYPODGERS 186
QY 182 KPQKXGNDVQDCIQWVTDIQTAVRTNSTFVQALVEHYKEGCDRLGPGMADICKNYISQ 241
DB 187 QPQKXGNDVQDCIQWVTDIQTAVRTNSTFVQALVEHYKEGCDRLGPGMADICKNYISQ 246
QY 242 YSEIATOMMHMOPKEICLVGFCDEVKEMQTLVPAKVASKNVLPALTEVPIKKEHV 301
DB 247 YSEVGVOMLHMOPKEICLVGFCDEVKEMQTLVPAKVASKNVLPALTEVPIKKEHV 306
QY 302 PAKSDVYCEVCEFLVETKLTIDNNKTEKILIDAPDKMSKLPKSLSECCQEVDTYSS 361
DB 307 QAHNVILCTCQCFVNMKSEELIVNNAATELLVKGSLSNACALLPDPARTKCEVGTGFP 366
QY 362 ILSILLESVPELVCSMLHLCG-----TRLPAL 390
DB 367 LLDIFLHEVNPSTLCGVLICLCAPELVEALEQAPATVSAALLKEPTPKQAPQKQAL 426
QY 391 TVVHTOPKQKGFCEVCKLVGYLDNLEKNTKQKIIALKEGCSFLPDPYQKQDQFVA 450
DB 427 PAHVPRQKNGFCVCEVCKLVGYLDNLEKNTKQKIIALKEGCSFLPDPYQKQDQFVA 486
QY 451 EYEPVILITLVEMDPSFVCLKIGACPSAHKPELLGTEKCIWGPSTYSCONTEAQCNAVE 510
DB 487 EYEPVILITLVEMDPSFVCLKIGACPSAHKPELLGTEKCIWGPSTYSCONTEAQCNAVE 546
QY 511 HCKRHVWN 518
DB 547 HCKRHVWN 554
RESULT 2
QTSY70 PRELIMINARY; PRT; 512 AA.
AC QTSY70.
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RT "genetic and genomic tools for Xenopus research: The NIH Xenopus
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshikiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy U., Heiton E., Kelleman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smalins D.E., Schmech A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP [3]
RC SEQUENCE FROM N.A.
RC TISSUE=Whole;
RA Klein S., Strausberg R.,
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC054988; AAH54988.1; -
RW Hypothetical protein.
SQ SEQUENCE 512 AA; 57555 MW; 57CFA5E2093F6FB1 CRC64;
Query March 52.1%; Score 1437; DB 13; Length 512;
Best Local Similarity 46.6%; Pred. No. 6e-101;
Matches 240; Conservative 130; Mismatches 131; Indels 14; Gaps 7;
QY 5 LGALAGPVILGKECTRGSAVWCONVKTASDCGAVHCLQTWNKPRVKSLEPDICKQV 64
DB 11 LALVATPLFGTRGCKAGPEVCEVNRTRISQCAVYKQOSVWNPVYKSLPCDVCCKIT 70
QY 65 TAAAGMLKDNATEBEILIVLEKTCMLPKPNMSASCKEIVDSYLYPIILIKGEMSRPGE 124
DB 71 TVVGNMKNITQGEIKYLVNKCILFPDPCGLAATCKQEVDSYFVILNLQDELSTNGV 130
QY 125 VCSALNLCESLOKHAELNHOKELESNKIPELDMTEVAVPAPNATLILYPODGERSRQ 184
DB 131 LCSSIGLCTSLORHIALSLKQPKQILNNEIPVDASKLVYPTVNPQLLYPOD--KTPKE 188
QY 185 PKQKXGNDVQDCIQWVTDIQTAVRTNSTFVQALVEHYKEGCDRLGPGMADICKNYISQSE 244
DB 189 PK-TGDIQNDCTTLLSDVQDALRSNSFSKTLVDHPQECNLDPPTAEMCKSYINQYAD 247
QY 245 IAIOMMHMOPKEICLVGFCDEVKEMQTLVPAKVASKNVLPALTEVPIKKEHVPAK 304
DB 248 IAIQVLMQPKQKQCGMAGFCDOEKSTPLQNIIPA---KSLIVAVITE---NPLP-E 298
QY 305 SDVYCEVCEFLVETKLTIDNNKTEKILIDAPDKMSKLPKSLSECCQEVDTYSSILS 364
DB 307 QAHNVILCTCQCFVNMKSEELIVNNAATELLVKGSLSNACALLPDPARTKCEVGTGFP 366
QY 362 ILSILLESVPELVCSMLHLCG-----TRLPAL 390
DB 367 LLDIFLHEVNPSTLCGVLICLCAPELVEALEQAPATVSAALLKEPTPKQAPQKQAL 426
QY 391 TVVHTOPKQKGFCEVCKLVGYLDNLEKNTKQKIIALKEGCSFLPDPYQKQDQFVA 450
DB 427 PAHVPRQKNGFCVCEVCKLVGYLDNLEKNTKQKIIALKEGCSFLPDPYQKQDQFVA 486
QY 451 EYEPVILITLVEMDPSFVCLKIGACPSAHKPELLGTEKCIWGPSTYSCONTEAQCNAVE 510
DB 487 EYEPVILITLVEMDPSFVCLKIGACPSAHKPELLGTEKCIWGPSTYSCONTEAQCNAVE 546
QY 511 HCKRHVWN 518
DB 547 HCKRHVWN 554
RESULT 3
QTSY74
QY 424 QEILALAEKGSFLPDPYQKQDQFVAEYEPVILITLVEMDPSFVCLKIGACPSAHKPL 483
DB 419 SRIDFLNRLINFLPDMQKQCSALIKVEYPLQLLEALDPSFLCLKLHLC-QGETVL 477
QY 484 LGTEKCIWGPSTYSCONTEAQCNAVEHCKRHVWN 518
DB 478 LGTEKCIWGPSTYSCONTEAQCNAVEHCKRHVWN 512

ID	SEQUENCE	PRELIMINARY:	PRT:	520 AA.
AC	Q8UVZ4			
DT	01-MAR-2002 (TReMBLrel. 20, Created)			
DT	01-MAR-2003 (TReMBLrel. 20, last sequence update)			
DT	01-OCT-2003 (TReMBLrel. 25, last annotation update)			
DE	Lysosomal cofactor/neurotrophic factor prosaposin.			
OS	Brachydanio rerio (zebrafish) (Danio rerio).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;			
OC	Cyprinidae; Danio.			
OX	NCBI_TaxID=7955;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Seo H.-C., Lie O., Fjose A., O'Brien J.S., Kishimoto Y.;			
RT	"Cloning, expression and promoter analysis of zebrafish prosaposin.";			
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF276996; AAL54381.1; -			
DR	GO; GO:0005764; C:Lysosome; IEA.			
DR	GO; GO:0006665; P:sphingolipid metabolism; IEA.			
DR	InterPro; IPR003119; SAPA.			
DR	InterPro; IPR007856; SAPB_1.			
DR	InterPro; IPR008138; SAPB_2.			
DR	InterPro; IPR008140; SAPB_sub.			
DR	InterPro; IPR008373; SAPosin.			
DR	InterPro; IPR008139; SAPosinB.			
DR	Pfam; PF02199; SAPA; 2.			
DR	Pfam; PF05184; SAPB_1; 3.			
DR	Pfam; PF03489; SAPB_2; 4.			
DR	PRINTS; PRO1797; SAPOSIN.			
DR	ProDom; PD001732; SAPB_sub; 3.			
DR	SMART; SM00162; SAPA; 2.			
DR	SMART; SM00118; SAPB; 4.			
QO	SEQUENCE 520 AA; 57451 MW; F95620F84BA41CB5 CRC64;			

Query Match	51.2%	Score 1412	DB 13	Length 520
Best Local Similarity	47.0%	Pred. No. 4.9e-99		
Matches 246	Conservative 120	Mismatches 139	Indels 18	Gaps 8

Qy	4	LLGALAAPRYLGTKECTGSAWVCNKNYASPCGAWKECLQTVNKKPVTYSJPCICIDV	63
Db	8	LVTYAVASPLGTBECARGPRYWCQNVYTAISLCAGVHOCQNVNKKPQMKYTPCULCEV	67
Qy	64	VTAAQDMUKDNAITEELLVYEKTCMDLPKPNMSASCKEYDSYLPVILDIKGMSPRG	123
Db	68	LVVVEPOLLKDNVTESELGYLEKAQOLIPDEGLAQCCKEYIDNYPEVWVGIIQGEILDRG	127
Qy	124	EVCASALNCSLOKHLAELNHQOCESNKIPELMTTEVAPAPMANIPILLYPODPRSKP	183
Db	128	VVCGAGLGLCVSQQALIA---KAQULTSNEIPQVLDANQKVSPLINIPOLLYEE---KRE	180
Qy	184	QPKQNGVDCDCIOMWTDIQTAVRTNSFTVQALVEHVHEECDRJPGMADICKNYISOYS	243
Db	181	TPKQKGVDCODCVTFISPTQDEARVNNSPFIINTLQAQVENQCETLJPGMSDMCKEYISOYG	240
Qy	244	EIALQMMHNPKEKICALVGR-DEVKEMPQMTLYPAK-VASKNYIPALELVEPL----K	297
Db	241	PLVFOQLMSPQKDCIARPGFCPTQKQSPKMEKLLPASIPAKNMFPAKKEKRPATMPA	300
Qy	298	KHEVPAKSDVVCVEVFLVKEVTKLIDNNKTEKEKILDAFDKSCSKLPKSISECCOBYDT	357
Db	301	KNLVAVRDSPOCALCEYWKIEIENNIQDPTSAELVQAVEKCNILIPSTLQAQCDLIET	360
Qy	358	YGSSTLSILBEVSEBELYCSMLHLCSG-TRLPAL-TVHTQPKDGGFCVYCKGLVGYLDR	415
Db	361	YGOALITDLVQADKRYTCSFALTCSGVSHVMQKHFAA---GGFCVDCMAARYVDG	417
Qy	416	NLEKNSYKOEILAALEKSGSFLPDEYQKCDQFVAIEBEVLLIEILVEVNDPFCVCLIGA	475
Db	418	ILEQVATQSEIEAVYKVSFLPDAVKBECNQLIQEYBELVQALLQTLDPFVCMKIGA	477
Qy	476	CPSAHKPLLTGTEKCIWGPSYWCNTETLTAQCAVHECKRHVN	518
Db	478	CPEAVORLLGNQCSWGPAYKCKNQTARCNALDHCRRHVS	520

RESULT	4
09DG82	
09DG82	PRELIMINARY; PRT, 522 AA.
AC	09DG82;
DT	01-MAR-2001 (TREMBLrel. 16, Created)
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE	Prosaposin.
PSAP	
OS	Brachydanio rerio (zebrafish) (Danio rerio).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC	Cyprinidae; Danio.
NCBI_taxid=7955;	
OX	

SQ SEQUENCE 522 AA; 57671 MW; D3C15A305725C1CD CRC64;

Query Match	48.6%	Score 1341	DB 13	Length 522
Best Local Similarly	45.7%	Pred. No. 1.2e-93		
Matches 240	Conservative 116	Mismatches 149	Indels 20	Gaps 9

[illegible]

Db 418 DGIIEONTQSEIEEFAVLKVCSPLEPYAKDCBNCNOLIREYELVQLLQTLDDPFCVKLT 477

Oy 474 GACSAHNPILGTCKLMGSESYWCNTETAAOCNAVECHKHVN 518
|||:::||||:||||:||||:||||:
Db 478 GACEAVORLEGLNQCSWGPAYWCKNVQTARCNALNHCRHWMS 522

RESULT 5

QCIC1 ID QCIC1 PRELIMINARY; PRT: 525 AA.

AC QCIC1;
DT 01-MAR-2003 (Tremblrel_23, Created)
DT 01-MAR-2003 (Tremblrel_23, Last sequence update)
DT 01-OCT-2003 (Tremblrel_25, Last annotation update)
DE Hypothetical saposin A-type domain/saposin type B containing protein.
GN 2310020A2IRIK.
OS Mus musculus (Mouse).
OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
XC NMMLTaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
EMBL AK028455; BAC25961.1; -;
DR MGD; MGJ:1924193; 2310020A2IRIK.
DR GO; GO:0005764; C-Lysosome; IEA.
DR GO; GO:0006655; P-sphingolipid metabolism; IEA.
DR InterPro; IPRO03119; Sapa.
DR InterPro; IPRO07856; SApB_1.
DR InterPro; IPRO08138; SApB_2.
DR InterPro; IPRO08140; SApB_sub.
DR InterPro; IPRO08373; Saposin.
DR InterPro; IPRO08139; SaposinB.
DR Pfam; PF02199; SAPA_2.
DR Pfam; PF05184; SApB_1;
DR Pfam; PF03489; SApB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SApB_sub; 3.
DR SMART; SMO0162; SAPA_2.
DR SMART; SMO0118; SAPB_4.
DR Hypothetical protein.
SQ SEQUENCE 525 AA; 57350 MW; 374F6050CD4D223 CRC64;

Query Match 39.5%; Score 1089.5; DB 11; Length 525;
Beet Local Similarity 40.0%; Pred. No. 1.66-74;
Matches 211; Conservative 108; Mismatches 172; Indels 37; Gaps 9;

Oy 2 ASLIGALLAGVYLKCEKTRGSAYWCNVKTASCCGAVKCLQTVNKKPTVKSIPCDICK 61
: ||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 12 SGLIGAAPRASPISTPRCAKSGSEVCCDLAAACRAVRHCOSAVMNKPTVKSIVCG 71
|||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Oy 62 DVVTAGDMKDNAVTEEEILVYLEKTCMDLPKPMSASCSEIYDSYLPVILDIKEMSR 121
|||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 72 DVAAAAANGVNGPATGESDILLTSWKTCENLPQSSESKKCMYNVNNNSAAVLSMISAQERT 131
|||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Oy 122 P-GEVSCALNLCSLQGHIALNLHQQLLESNKIPELDMEVVAFPFMANIPLLLYPDGPR 180
|||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 132 DLASVCPALFLCPSLRGHLAETTSEBPILTQE-----DAEVVAAPFLSNGALSHPESQM- 185
|||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Oy 181 SKQPQXNDGVCCDCTQMTVDITQAVRINSTFYQALVEHYKEECDDLGGMADICKNYIS 240
|||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 186 -----EGAVCHDCQVLISLLDPALESNTTLAEVTQN---QCQSWGGLAALCENTYIH 235
|||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Oy 241 QYSEIAIQMMHMHPKEICALVGFCDEVKEMPQTLVPARKVASKNVTPALEVPEPKKE 300
: |||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

Db      236 RQFVPAKQTLQGLPQEQVCRKGGCFEERSAHL-----TRVAADVGVSLEMEHP-RTNE 28
Qy      301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFMCCKLPKSLSECOEVDVTGS 36
Db      290 LOMQGLTCDVCLNLVQELDKMLVTNSTEALISHTLERCVVPEPLVQCCITLVDTISP 34
Qy      361 SLISLLEEVSPPELVCSMLHLCSGTR-----LPALTAVYTOPKDGCEVCKKL 40
Db      350 ELVQ-LMSKVTPEKVCETIKLKGSGRRARSTSRVAATPSP-L-VDEENQGSFCQCKRL 40
Qy      410 VGYLDRLNKNSTYKQEILALAEKCSFLPDPEYQKQDFVLEYEPVLEILVWMDSPFV 46
Db      407 LQMSQNDLHKSTYRDIILNAFKGCGRIPLPYVWQCNRFVAEYEPVLESIKFMANPDL 46
Qy      470 CLKTGACGSAHKPILGTGECIKWGSYVCONETTAQCAVNEHCGRHW 517
Db      467 CKKGACGCGKPTPLGLTDQCVGSPFCKSPPEAEKMALEHCQRLVW 514

RESULT 6
Q8BJV5 PRELIMINARY; PRT; 449 AA.
ID Q8BJV5
AC Q8BJV5;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical saposin A-type domain/saposin type B containing protein.
GN 2310020A2IRIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eye;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the Riken Genome Exploration Research Group Phase I & II Team,
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK078699; BAC37363.1; -.
DR MGD; MGI:1924193; 2310020A2IRIK.
DR GO; GO:0005764; C.Lysosome; IEA.
DR GO; GO:0006655; P.sphingolipid metabolism; IEA.
DR InterPro; IPR003119; SAPA.
DR InterPro; IPR007856; SAPB_1.
DR InterPro; IPR008138; SAPB_2.
DR InterPro; IPR008140; SAPB sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA_1.
DR Pfam; PF03484; SAPB_1; 3.
DR Pfam; PF03489; SAPB_2; 3.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SAPB sub; 2.
DR SMART; SM00162; SAPA_1.
DR SMART; SM00118; SAPB_3.
DR Hypothetical protein.
KW SEQUENCE 449 AA; 48705 MW; 377CC42A475B292F CRC64;

Query Match 28.1%; Score 776; DB 11; Length 449;
Best Local Similarity 35.9%; Pred. No. 9,4e-51;
Matches 165; Conservative 90; Mismatches 160; Indels 44; Gaps 12;

Qy      2 ASLIGALLAPVGLGKECTGSAVQCNVKTASDCAVKKQLQTVWKKPTVKSLPDICK 61
Db      12 SGLIGARRAPISVPRCCARGSEWQDDLOAAKCAVRHCGSAVWNNKPTVKSIPGVCQ 71
Qy      62 DVYTAADMKADNATEEELIVLEKTCMDLPRKMSASCEIYDSVLPVILDIKGMER 121
Db      72 DVAAAGNGVNPAGTISDILTSYMKTCENLPSESSAKCKMAYNNNSAAVLSGAGQT 131

```

Qy	122	P-GEVSAIYNICE5IOKHLAEINHQ4UESNNKIPELDMTEVVAPEMANIPBLIYPODGR	160
Db	132	DLASVCTALTYLCEBPIORHIAETTSERPLQOE-----DANEVMAPIFISNGALSFHPQMP-	185
Qy	181	SKPCKMDNGVQDCDLOMTWTDIQTARTNSTFWQALVEHKEBCDRLGPGMADICKNYIS	240
Db	186	-----EGAVCHDCVQILSILODPALESNLTIAEVTYON--OCQSGPELALACEYIYH	235
Qy	241	QYSEIAIOMMMHOREKICALVGFCEDEVEMPMQTLVPAKVASKNYIPALELVEPIKHE	300
Db	236	RQFVPAKQTLQGLPQOEVCGRKSGFCGRESAHML-----TRVAVDGVPSLMEMMP-RTNE	289
Qy	301	VPAKSDVATCEVCEFLVKEVTKLIDNNKTEKEIILDAFDMKSKLPKLSBEOEAVDTYGS	360
Db	290	LQMVGQTCDCVCTALNVEQBELDKMLVTNSTEALISHTLERCTAVPEPLVQCCITPLVDTYP	349
Qy	361	SILSLILBEVSELYCSMLHLCSGTR-----PALTYHVATQPKDGGCEVQCKKL	409
Db	350	ELVQ-LMSKVPEKXCETIKLQGSRRPARISRAVATTPSLP--VDENQSGFCQCGCKRL	406
Qy	410	VGYLDNRLEKNSTKQEIILALEKGC--SEPLPYQKQD	446
Db	407	LGH-----VFPESGPEQHAG-HSECIQRWLPPDVALCD	440

RESULT 7

ID Q8C1N0 PRELIMINARY; PRT; 402 AA

AC OBC1N0; 2003 (TEMBLrel. 23, Created)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Hypothetical saposin A-type domain/saposin type B containing
 DE protein.
 GN 2310020A21RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE= Tongue;
 RX MEDLINE=22354683; PubMed=1246681;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK009408; BAC2528.1; -;
 DR MGD; MGI:1924193; 2310020A21RIK.
 DR GO; GO:0005764; C:lysosome; IEA.
 DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
 DR InterPro; IPR003119; Sapa.
 DR InterPro; IPR007856; Sapa.1.
 DR InterPro; IPR008138; Sapa.2.
 DR InterPro; IPR008140; Sapa.sub.
 DR InterPro; IPR008373; Saposin.
 DR InterPro; IPR008139; SaposinB.
 DR Pfam; PF02199; SAPA, 1.
 DR Pfam; PF05184; Saps_1; 3.
 DR Pfam; PF03489; Saps_2; 3.
 DR PRINTS; PRO1797; SAPOSIN.
 DR ProDom; PD001732; Saps sub; 2.
 DR SMART; SM00162; SAPA, 1.
 DR SMART; SM00118; SApB; 3.
 DR Hypothetical protein.
 QO SEQUENCE 402 AA; 44420 MW; E90017CBFA017ED6 CRC64;

Query Match	28.0%;	Score 772;	DB 11;	Length 402;
Best Local Similarity	38.1%;	Pred. No. 1.6e-50;		
Matches 154;	Conservative 85;	Mismatches 129;	Indels 36;	Gaps 8

Qy	125	VCALNCSLQKHIAELNHOQLESNNKIPELDMTEVYAPPMANIPILLYYQDQSPKQP	184
Dp	13	VCYALTLCEBPLOHIAETTSERPLTQ-----DANEVAPFLUSNGALSFFHSQMP-----	62
Qy	185	PKDNGDVDCDIOMVTDIQTAVRTSTFVOALVEHVKESCRRLPGMADICKNYISQYSR	244
Dp	63	---EGAVCHDCQVLTISLQDNLJESUPLTABLYQV---QQCSMGFLAALCENYTHHQFV	116
Qy	245	IAIOMMHHQPYEICALVCFDEVKEMPMQTLVPKAVASKNVLPALBIVEBPKHVPAPK	304
Dp	117	PAKQTLQGPBPQVCKGFCGERESAHVLP-----TRVAALVGVPSLLEMMP-RTNELQMQ	170
Qy	305	SDVYCCVCEFLVKEVYKTLIDNNKTEKELIDAFDKMCSLTPKLSBEOCEVVDYTGSSILS	366
Dp	171	LGITCOVCNLNVOELDKMLVYNSTEALISHTLERCTVAPPEPLVQQCTTLDVYISPELVQ	230
Qy	365	ILLEEVSPELVOSMHLTCSGTR-----LPALTVAHTOPKDGFCVCEKCLVGYL	413
Dp	231	-LMSKYLTPKVCETIKLGOSKRRARSISRAVATPSP--VDENQGSFOCGCRRLQMS	287
Qy	414	DRNLKRNSTKQBIILAALEKGSFLPDYQKQDQFVAEYEBVLLIILVEVMDPSFVCLKI	473
Dp	288	SQMLDKHSTKRDILNAAFKGCRILPLPYVMQCNRFVAEYEBVLLIESLKFMNPTDLCKM	347
Qy	474	GAQPSAHRKLLGTCKCIWMPSTWQONTETAQNCNAVEHCKRHV	517
Dp	348	GAGHGKPTLLGTQCVMPGSPFWCKSPRAAEKCNALHICQSLV	391

RESULT 8

ID Q9Y125 PRELIMINARY; PRT; 953 AA.

DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE SAV-R protein.
GN SAV-R OR BCIDNA:GH08312 OR CG12070.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abil J.F., Aghayani A., An H.-J., Andrews-pfankuch C., Baldwin D.,
RA Balcer R.M., Basu A., Baxendale J., Bayraktiroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bock J., Brokstein P., Brottier P.,
RA Burts K.C., Busch D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cleyer S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fessler C., Gabriellian A.E., Garg N.S., Galbraith W.M., Glasser K.,
RA Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatalai B., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Kechum K.A.,
RA Kammel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Liao X., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paley J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spter E., Spredling A.C., Stapleton M., Strong R., Sun E.,
 RA Switkes R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 RN
 RP SEQUENCE FROM N.A.
 RA Rudin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
 RA Agapayni A., Arcaina T.T., Baxter E., Blazey R.G., Butenhof C.,
 RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
 RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
 RA Houston K.A., Hummachi S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,
 RA Park S., Seguelia A., Sethi H., Smit E., Switkes R.R., Weinburg T.,
 RA Ceiniker S.E.,
 RT "Full length *Drosophila melanogaster* cDNA sequence,"
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AEO03775; AAF57097.1; -
 DR EMBL: AF145647; AAD38622.1; -
 DR FlyBase: FBgn0000416; Sap-t.
 DR GO: GO:0005764; C:lysosome; IEA.
 DR GO: GO:0006665; P:sphingolipid metabolism; IEA.
 DR InterPro: IPR003119; SApA.
 DR InterPro: IPR007856; SApB.1.
 DR InterPro: IPR008138; SApB.2.
 DR InterPro: IPR008140; SApB_sub.
 DR InterPro: IPR008373; Saposin.
 DR InterPro: IPR008139; Saposinb.
 DR Pfam: PF02189; SApA.1.
 DR Pfam: PF05184; SApB.1; 6.
 DR Pfam: PF03489; SApB.2; 7.
 DR PRINTS: PRO1797; SAPOSIN.
 DR ProDom: PD001732; SApB_sub; 6.
 DR SMART: SM00162; SApA.1.
 DR SMART: SM00118; SApB.1.
 DR SMART: SM00118; SApB.7.
 SO SEQUENCE 953 AA; 105962 MW; D6CFED3E9D1502A8 CRC64;
 Query Match 19.9%; Score 549.5; DB 5; Length 953;
 Best Local Similarity 23.2%; Pred. No. 4e-33;
 Matches 144; Conservative 105; Mismatches 212; Indels 161; Gaps 18;
 QY 8 ALAGVGLGKECTRGSAVWCQNVKTASDCGAVKICLTQVWNR---PTVSKSPDICGV 64
 DB 20 AATATLLSSKCTWGPSTWCGFNSKCRATRHICITQVWETQKVPVDTISICTCKDV 79
 QY 65 TAAGDMLKDNATTEETIILVLEKTCQDMLPKPMMSASCKEIVDSYLPVLDTIKGMSRGE 124
 DB 80 TQARQQLSNQTEBELKEVFEGSKCLIFIKTQKECTKIVADDFELVEALASQMN-PDQ 138
 QY 125 VCSALNLCES-----LQKH-----LAELN 143
 DB 139 VCSVAGLCNSARIDELYNKGIQAGLDGTQVQNEDESSSETELAMQPNQSCGNCNLSRLM 198
 QY 144 HOKQLESKIKPELMTETV-----APFANITPLI----- 173
 DB 199 HSKRAADRRD---DMVETLMHMGSSLSFSDACANITVLTFFNDIYDHSKHLTTDAVCHV 255
 QY 174 -----YPODGRSKPOPK---DNQD-----VCQDCIQMTVDIGTAVRTNSTPFOALV 217
 DB 256 SGVCASHYHOHEEKQPOEALVMDAGDDIPCELCQGLVGHLDVAVANTTEFEFQVME 315
 QY 218 EHVKEECDRIGPGMADICKNYISQYSRTAIC-MMMHMQPREKICALVGFDE-----VKEM 271
 DB 316 GFCKQ-----SKGFQECISIVQYHVIETLVSKLDNGACMIGICQKNSASMSKOV 370
 QY 272 PMQTLVPAKVASIKVLPALTEL---VEPIKKE-----VP----- 302

DB 371 PIMPLP-----VIEPAQKITTEKLEKKEKQQLASPEKPSQOEILDMQPIHLMG 423
 QY 303 -----AKSDYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCSTLPSISECCOEYV 355
 DB 424 AANPAGVGEGLTCLCEYMHFIQETLAPSTODEIKHTVENICATLPSGVAGQCRNFV 483
 QY 356 DTYGSSILILEEVSPELVCSMLHLCSTGRPLATYHTQP-----KDGFEVCCKLU 409
 DB 484 EMTGDAVIALVQGLNRPDVCPLMQCPKPLPKXEDVEVNPQASDEQDPPTCPLCPA 543
 QY 410 VGYIDRLNEKXSTKOEILALKEKCSFLPDPYQCCQDFVAEYEPVLIETLVEWMPSPV 469
 DB 544 VEOQMIRDNKSKDKNTKVTNLGCLSHLPRIKEECVDFPNTYSNLIIMLIDFDPQEI 603
 QY 470 CLKTGACPSAKKPLTGKECIM 491
 DB 604 CVQKLCPK-----TTVALW 618

RESULT 9

Q8N7T4 PRELIMINARY; PRT; 241 AA.
 AC Q8N7T4
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein FLJ40379.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 RN NCBI_TaxId=9606;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
 RA Arita M., Masahino K., Yunki H., Hara H., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Magatsuma M.,
 RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
 RA Sugano S., Nagahori K., Maehuo Y., Nagai K., Teigai T.,
 RT "NEO human cDNA sequencing project,"
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK097698; BAC05143.1; -
 DR GO: GO:0005764; C:lysosome; IEA.
 DR GO: GO:0006665; P:sphingolipid metabolism; IEA.
 DR InterPro: IPR003119; SApA.
 DR InterPro: IPR007856; SApB.1.
 DR InterPro: IPR008138; SApB.2.
 DR InterPro: IPR008140; SApB_sub.
 DR InterPro: IPR008373; Saposin.
 DR InterPro: IPR008139; Saposinb.
 DR Pfam: PF02189; SApA.1.
 DR Pfam: PF05184; SApB.1; 2.
 DR Pfam: PF03489; SApB.2; 2.
 DR PRINTS: PRO1797; SAPOSIN.
 DR ProDom: PD001732; SApB_sub; 2.
 DR SMART: SM00162; SApA.1.
 DR SMART: SM00118; SApB.2.
 KW Hypothetical protein.
 SO SEQUENCE 241 AA; 26719 MW; BE29EEAF6EEB6AB CRC64;

Query Match 19.8%; Score 545; DB 4; Length 241;
 Best Local Similarity 44.6%; Pred. No. 1.6e-33;
 Matches 107; Conservative 45; Mismatches 76; Indels 12; Gaps 4;

QY 287 IPALTEVPIKKEHVPAKSVYCEVFLVKEVTKLIDNNKTEKEILDAFDKMCSTLPS 346
 DB 4 VPSLELGLPKQSEMQKQGVTCVCMNVQKIDHLMKNSSELMITHALREVCSVMPAS 63
 QY 347 LSEECQVVDVTSSSILILEEVSPELVCSMLHLCSTGRPLATYHTQP-----VTQPK--- 398
 DB 64 ITRKCIIVDTYSPSLVQ-LVAKITPEKVCKFRILC-GNRRRAVRAVDAVATVPSPEWDA 121

QY 399 --DGGFCEVCKLVGLYLDNLEKSTKOBLLAALFKGCSFLPDYQKODQFVAEYEPVL 456
DB 122 ENQSGFCNGCKRLTVSSHLESKSTKRIILVAFKGCILPLPYMIQCKHVTYQEPVL 181
QY 457 IEILVEVNDPSPVCLIKIGACPSAHRPLLGTEKCIQNGPSYWCNTETAACNAVEHCKRHV 516
DB 182 IESLMDMNDPVAVCKKVGACHPRPFLICTDQCALGSPSWCSQBAKLCNAVHCQKHV 241

RESULT 10
ID P79254 PRELIMINARY; PRT; 121 AA.
AC P79254
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Sulfated glycoprotein-1/SGP-1 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96249303; PubMed=8848570;
RA Spencer T.E., Graf G.H., Bazer F.W.;
RT "Sulfated glycoprotein-1 (SGP-1) expression in ovine endometrium
during the oestrous cycle and early pregnancy.";
RL Reprod. Fertil. Dev. 7:1053-1060(1995).
DR EMBL; S82555; AAD14405.1; -.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR008138; SAPB_2.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF03489; SAPB_2; 1.
DR PRINTS; PR01797; SAFOSIN.
DR SMART; SM00118; SAPB; 1.
FT NON_TER 1
SQ SEQUENCE 121 AA; 13604 MW; 4F0F5A6EB83D0C9A CRC64;

Query Match 18.7%; Score 517; DB 6; Length 121;
Best Local Similarity 77.7%; Pred. No. 9e-32;
Matches 94; Conservative 18; Mismatches 9; Indels 0; Gaps 0;

QY 199 VTDIOTAVRTNSTFYQALVEHVEKCDRLGPGMADICKNYISQYSEIALQMMHNPKEI 258
DB 1 VTDIOTAVRTNSTFYQALVEHVEKCDRLGPGMADICKNYISQYSEIALQMMHNPKEI 60

QY 259 CALVFCDEVEKEMPQTLVPAFVASKNVIPLALEVEPIKHEVPAKSDVYCEVCEFLVKE 318
DB 61 SALVFCDEVEKEMPQTLVPAFVASKNVIPLALEVEPIKHEVPAKSDVYCEVCEFLVKE 120

QY 319 V 319
DB 121 V 121

RESULT 11
ID O15997 PRELIMINARY; PRT; 965 AA.
AC O15997;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Bmp109.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=98278844; PubMed=9611271;
RA Tambunan J., Chang P.-K., Li H., Natori M.;
RT "Molecular cloning of a cDNA encoding a silkworm protein which
contains the conserved BH regions of Bcl-2 family proteins.";
RL Gene 212:287-293(1998).
DR EMBL; AB008449; BAA23126.1; -.
DR PIR; T00207; T00207.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; SAPA.
DR InterPro; IPR007856; SAPB_1.
DR InterPro; IPR008138; SAPB_2.
DR InterPro; IPR008140; SAPB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SAPB_1; 6.
DR Pfam; PF03489; SAPB_2; 7.
DR PRINTS; PR01797; SAFOSIN.
DR Prodom; PD001732; SAPB_sub; 6.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 7.
SQ SEQUENCE 965 AA; 108825 MW; FA1A7BEETF626078 CRC64;

Query Match 18.5%; Score 511; DB 5; Length 965;
Best Local Similarity 24.6%; Pred. No. 3.5e-30;
Matches 139; Conservative 104; Mismatches 229; Indels 94; Gaps 18;

QY 13 VLGLKECTRGSAVWCQNTKATSDCGAVGCHQOTWNNK---PTVASLPEDICKDVTAAGD 69
DB 170 LLGKRCRTWGPSVCSNSTGRCNATPHCINRWVSKTFEDNDNTOICLIDWYKQARD 229

QY 70 MKDNATSEELVLEKTCMLPKPNMSASKEIYDSTYLPVILDIKEMSRGEGVCSAL 129
DB 230 QLOSNETODEIKEYEVGSCKLIPKFVAGCMKLADBEFVBLIETLASEMN-PQAVCSVA 288

QY 130 NLCEB--LQKILAEINHKQLES-----NKIPELDMTEVVAAPFA----- 167
DB 289 GLCNNAKIDRLLVNDNAQRELRAQCYNCQKTGVVKKFDETKYEDFLVGLLYQVGRNDS 348

QY 168 ---NIPLLVY-----QDG-----PRSKPQPKNGD 190
DB 349 LSDSCSMILFKYENILEAVKDLNPEGICHVSGQSYKFNHDEFTPEQWQVSARD 408

QY 191 V-CODCIQMTVDIOTAVRTNSTFYQALVEHVEKCDRLGPGMADICKNYISQYSEIALQ 249
DB 409 VPCEFCDEGLVGNLRVLVAANTTELE-FYKVLQGLCKQTKG-FKDECLHLAEQYVVIYNF 466

QY 250 MM-HMOPKEICALVFCDEVEKEMPQTLV-----PAKYA-----SKNVIP 288
DB 467 LVSDLKPAETCKMIGICNLTSAPIPSLVARELVKVPKILGAESEKIAVPLAKOMEP 526

QY 289 ALLEVE--PIKHEVRA-KSDVYCEVCEFLVKEVYKLLDNKTEKEIILDAPDKMSKLPK 345
DB 527 ASAAVSVPLERMFLAQPQSKAACAFQYFLHYLOVLSDRTEBKVAQAQACDALPD 586

QY 346 SLSECOEVDVTGSSILSLLEEVSPVLVCSMLHCSGTRLPALTVAHTVOPKDGFCFV 405
DB 587 ALNGECKEFVQYGSAAVIALVQEIDPASVCPALDICTQTS-EIRRVVNSEKSN--CPL 643

QY 406 CKLVGLYLDNLEKSTKOBLLAALFKGCSFLPDYQKODQFVAEYEPVLIELIVEVND 465
DB 644 CLFAVEQLESVLYKNRSESENIKALDGLCTRLSQDQSECIDFVDTYSQLVEMLVADMN 703

QY 466 PSFVCLRTKIGAC-PSAHRPLLGTEKCI 490
DB 704 AKEICVPLKCRDQLHDLKLTHTSSI 729

RESULT 12
ID Q8IMH4 PRELIMINARY; PRT; 876 AA.
AC Q8IMH4
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Bmp109.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]

AC 081M4;
 DT 01-MAR-2003 (Tremblrel. 23. Created)
 DT 01-JUN-2003 (Tremblrel. 24. Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25. Last annotation update)
 DE CG12070-PB.
 CN SAp-R.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NX NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=2019606; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazer V., Chao M., Pfeiffer B.D.,
 RA Abol J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Berens P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibeagwu C.,
 RA Jalali M., Kalish J., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Klamel B.E., Kodita C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodgett M., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 [2]
 RP SEQUENCE FROM N.A.
 RP Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Bantz C.J., An H., Baldwin D., Bantz J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,
 RA Dodson K.J., Dorsett V., Doup L.E., Doyle C., Drenek D., Fafian D.,
 RA Ferreira S., Frisbe E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibeagwu C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nuno J.,
 RA Pauley J.M., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svitek R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RP Maira S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochman S.E., Smith C.D.,

RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celniker S.E.,
 RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
 RA Harris N.L., Kronmiller B., Marshall B., Milburn G.H., Richter J.,
 RA Russo S., Seale S.M.J., Smith E., Shu S., Smutnack F.,
 RA Whitfield B.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
 RA Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RP FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RP FlyBase;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003775; AN14261.2;
 DR GO; GO:0005764; C:lysosome; IEA.
 DR GO; GO:0005489; F:electron transporter activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
 DR InterPro; IPR000345; CytC_heme_BS.
 DR InterPro; IPR007856; Stp1.
 DR InterPro; IPR008138; Stp2.
 DR InterPro; IPR008140; Stp3.
 DR InterPro; IPR008373; Stp4.
 DR InterPro; IPR008139; Saposin.
 DR Pfam; PF05184; Stp1_1.
 DR Pfam; PF03489; Stp2_1.
 DR PRINTS; PR01797; SAPOSIN.
 DR ProDom; PD001732; Stp sub. 2.
 DR SMART; SM00118; SAPB_7.
 DR PROSITE; PS00190; CYTOCHROME_C_1.
 SO SEQUENCE 876 AA; 97617 MW; 767E16D35ACF52DB CRC64;
 Query Match 15.7%; Score 432.5; DB 5; Length 876;
 Best Local Similarity 23.8%; Pred. No. 2.9e-24;
 Matches 117; Conservative 91; Mismatches 178; Indels 105; Gaps 17;
 QY 55 LPEDIDKVVYTAGMNLKDNATTEBILVLEKTCQDMLPKPMMSASCKEIVSYLPVLDI 114
 DB 208 IPELCEQVYKRLRVLVANTETETFEKQVMEGFC- QSKGFKDECSIVDQYHVHIVET 265
 QY 115 IKGMSRPGEVCSALNLC-----ESLQKHLAELNHQ 145
 DB 266 LVSKIDANG-ACCMGICQKNSASMKDVPIMPLVIEPAQVKTTEKLEKH-----EK 319
 QY 146 KQLESNKIPELDMTEVAPFMANIPLLLYPODGPRSKROP---KXNGDVCCDCTQWVTDI 202
 DB 320 KQLGASE-PKRSQGIID-----MQLPIDHMGANPGALVEGSELCTLCYMLHFI 370
 QY 203 CFAVNTSTFQALVENVKEE-CDRLPGNADICTKNTISQYSEIATQMMNH-MQPKICA 260
 DB 371 QETLTPESTIDE--IKHVENICAKLPBGVAGQCRNFVEMVGDVIALVQGLNPRDVC 428
 QY 261 LVGFCDEVKEMPMQTLVPAKVASKNVIPALETVEPIKHEVPAKSDVYCVCEFLVKEVT 320
 DB 429 LMQWCP-KNLPKK-----EDVEVNPQPSADBDP-----PT-----CPLCFPAVDAQ 471
 QY 321 KLIDNKTKEKILAFDMKSKLPKXISECCQEVVDYTGSSISILLESVPALVCSMTH 380
 DB 472 MKIDNNSKDKIKVNLGSLSH-PNEIKECVDVFNVTYSELIDMLITDPKPEICVQLK 531
 QY 381 LCSGT-RLPALTVHVTQPKDG-----GF-----CEV 405
 DB 532 LCPRTTALMDLRISLDDVDGDKSSSEISINDIESLEELPPQALFDPGFPAAENCI 591
 QY 406 CKKVLGVLDNRNLKNSKQBIALALEKGSFLPDPYQKODGPFVAYEYPLVLEIIVEND 465
 DB 592 CEELVKTLEKRMGKHPRDSIKHILESSCDRMKPKPNTTCHVYIDYGGKIDLLKEND 651
 QY 466 PSFVCLKIGAC 476

Db 652 PKLICTELGMC 662

RESULT 13

Q95X02 PRELIMINARY; PRT; 458 AA.
 AC Q95X02;
 DT 01-DEC-2001 (TREMBLrel. 19, Created).
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Naegleria fowleri.
 GN NP-B.
 OS Naegleria fowleri.
 OC Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.
 OX NCBI_Taxid=5763;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nickel R., Leipe M.,
 RT "Pore-forming peptides of Naegleria fowleri."
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF196309; AAL01158.1; -.
 DR GO; GO:0005764; C:lysosome; IEA.
 DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
 DR InterPro; IPR007856; SApB_1.
 DR InterPro; IPR008138; SApB_2.
 DR InterPro; IPR008140; SApB_sub.
 DR InterPro; IPR008139; SaposinB.
 DR Pfam; PF05184; SApB_1; 5.
 DR Pfam; PF03489; SApB_2; 5.
 DR PRINTS; PR01797; SAPOSIN.
 DR Prodom; PD001732; SApB_sub; 4.
 DR SMART; SM00118; SApB; 5.
 FT NON_TER
 SQ SEQUENCE 458 AA; 50452 MW; 433F60CDE64C904 CRC64;

Query Match 12.5%; Score 343.5; DB 5; Length 458;

Best Local Similarity 25.6%; Pred. No. 7.4e-18; Matches 115; Conservative 69; Mismatches 186; Indels 79; Gaps 16;

QY 36 CGAVHGC---LQTVNKKPTVKSIPDCIDKDVTAAGMILKONATEEELVLYLEKTCMDLP 92
 DB 69 CSQVTKCNGSAMAAVAAPRAENSGICNMQLVTVQENNVESNDITMTLEKLEQYCVSVP 128
 QY 93 KPNMSASKEIYDSYLPVILDIKGMSPG-EVCSALNLCESLQKHLAELNHQQLBSN 151
 DB 129 -GQYSALCTYAVBOYLPIFIHQV--EKOPPALITICQDVHLCSAQ----- 170
 QY 152 KIPELDMTEVVAPEMANIPLLLYPODGRSKPOPKDNGVQDCIQMTDIOQAVRTNST 211
 DB 171 -----AAPVV-----QQQQAALCPICRAVGFLETKI--NNV 201
 QY 212 FVQALVHVKRECDRLGPMADICKNYISQYSEIAIQMMHMQPEIKCALVGFCDVEKEM 271
 DB 202 DVNAVKQQLERACSF--QVPDCCQIVDKAQLA-QDLQTEDAQITCSTV-----VVC 252
 QY 272 PMQITVPAKVASKNVIPALETVEPIKGEVPAKSDVYCEVGFYKEVTKLIDNNKTEKE 331
 DB 253 PKQGVV-----TFNPFRKF--LEAKDSKYCTCQITQITYLEDLIVSDITVNE 297
 QY 332 ILDAFDKMSCKLPKSLSEECQEVVDYTGSSILSILEEVSPELVCSMLHCSGTRLPALT 391
 DB 298 IIKLADAGCARL-GALIESLCKKFFVPLAVDELKLLLEKLPQKVCSTLKMCDAAELIKLA 356
 QY 352 VHAVTPKDGGFCEVCKKLVGLDNLNLEKNSIKQELIALLEKGC-SFLDPDYKQCDQFYA 450
 DB 357 L-APQADGTCMLACEYVISVADNMLANNTOQSVKNTLDKVCQGFVPSIYOSOCIALVN 415
 QY 451 EYEPVLIETL-VEVWDPSFVCLKIGACPS 478
 DB 416 QYEAQLIQLFESKVFNPQTVCKAIGVCS 444

RESULT 14

Q9BKML PRELIMINARY; PRT; 484 AA.
 AC Q9BKML;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Naegleria fowleri.
 GN PRONP-B.
 OS Naegleria fowleri.
 OC Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.
 OX NCBI_Taxid=5763;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nickel R., Benkert C., Jacobs T., Marti T., Marciano-Cabral F.,
 RA Leipe M.,
 RT "Pore-forming peptides of Naegleria fowleri."
 RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF154047; AAK21659.1; -.
 DR GO; GO:0005764; C:lysosome; IEA.
 DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
 DR InterPro; IPR007856; SApB_1.
 DR InterPro; IPR008138; SApB_2.
 DR InterPro; IPR008140; SApB_sub.
 DR InterPro; IPR008139; SaposinB.
 DR Pfam; PF05184; SApB_1; 5.
 DR Pfam; PF03489; SApB_2; 5.
 DR PRINTS; PR01797; SAPOSIN.
 DR Prodom; PD001732; SApB_sub; 4.
 DR SMART; SM00118; SApB; 5.
 SQ SEQUENCE 484 AA; 53237 MW; 63DA3A8E7CA578B0 CRC64;

Query Match 12.5%; Score 343.5; DB 5; Length 484;

Best Local Similarity 25.6%; Pred. No. 8e-18; Matches 115; Conservative 69; Mismatches 186; Indels 79; Gaps 16;

QY 36 CGAVHGC---LQTVNKKPTVKSIPDCIDKDVTAAGMILKONATEEELVLYLEKTCMDLP 92
 DB 95 CSQVTKCNGSAMAAVAAPRAENSGICNMQLVTVQENNVESNDITMTLEKLEQYCVSVP 154
 QY 93 KPNMSASKEIYDSYLPVILDIKGMSPG-EVCSALNLCESLQKHLAELNHQQLBSN 151
 DB 155 -GQYSALCTYAVBOYLPIFIHQV--EKOPPALITICQDVHLCSAQ----- 196
 QY 152 KIPELDMTEVVAPEMANIPLLLYPODGRSKPOPKDNGVQDCIQMTDIOQAVRTNST 211
 DB 197 -----AAPVV-----QQQQAALCPICRAVGFLETKI--NNV 227
 QY 212 FVQALVHVKRECDRLGPMADICKNYISQYSEIAIQMMHMQPEIKCALVGFCDVEKEM 271
 DB 228 DVNAVKQQLERACSF--QVPDCCQIVDKAQLA-QDLQTEDAQITCSTV-----VVC 278
 QY 272 PMQITVPAKVASKNVIPALETVEPIKGEVPAKSDVYCEVGFYKEVTKLIDNNKTEKE 331
 DB 279 PKQGVV-----TFNPFRKF--LEAKDSKYCTCQITQITYLEDLIVSDITVNE 323
 QY 332 ILDAFDKMSCKLPKSLSEECQEVVDYTGSSILSILEEVSPELVCSMLHCSGTRLPALT 391
 DB 334 IIKLADAGCARL-GALIESLCKKFFVPLAVDELKLLLEKLPQKVCSTLKMCDAAELIKLA 382
 QY 352 VHAVTPKDGGFCEVCKKLVGLDNLNLEKNSIKQELIALLEKGC-SFLDPDYKQCDQFYA 450
 DB 383 L-APQADGTCMLACEYVISVADNMLANNTOQSVKNTLDKVCQGFVPSIYOSOCIALVN 441
 QY 451 EYEPVLIETL-VEVWDPSFVCLKIGACPS 478
 DB 442 QYEAQLIQLFESKVFNPQTVCKAIGVCS 470

RESULT 15

Q35489 PRELIMINARY; PRT; 378 AA.
 ID Q35489

```

AC 035489;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Surfactant protein-B.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Cranialata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RA Bingle C.D., Yuan H.T., Gowan S.;
RT "Guinea pig surfactant protein-B is differentially polyadenylated.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF033190; AAB87086.1; -.
DR HSSP; P07988; 1DFW.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0007585; P:respiratory gaseous exchange; IEA.
DR GO; GO:0006655; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; Sapa.1.
DR InterPro; IPR008138; Sapa.2.
DR InterPro; IPR008140; Sapa_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008137; SaposinB.
DR Pfam; PF02199; SAPA; 1.
DR Pfam; PF05184; Sapa_1; 2.
DR Pfam; PF03489; Sapa_2; 3.
DR PRINTS; PR01797; Saposin.
DR ProDom; PD001732; Sapa_sub; 1.
DR ProDom; PD008002; Surfactant_B; 1.
DR SMART; SM00162; SAPA; 1.
DR SMART; SM00118; Sapa; 3.
SQ
SEQUENCE 378 AA; 41677 MW; E3DA2E237ED401DA CRC64;

Query Match      12.4%; Score 342.5; DB 11; Length 378;
Best Local Similarity 22.6%; Pred. No. 76-18;
Matches 111; Conservative 67; Mismatches 174; Indels 139; Gaps 15;

QY 1 LASLIGALLAGFVLGKCTRGSAVWQNVKTAASDCGAVKHCLQTVWNNKPTVKSIPCDIC 60
DB 13 LPMLCGPGAAQTTSILACSGQPFKFCQSLBQALGCRALGLOEVMGHVAGADDL-CQEC 71
QY 61 KDVVTAAGDMLKDNATEEELVLEKTCDDMLPKPMNSASCKEIVDSYLPVLDITIKGEMS 120
DB 72 EDIVAILTKMAKEVLEFQTKIRFLERECDDLPLKLVPRCHSVLETFEPLVIDYFQSHIT 131
QY 121 RPEGVCSALNLCESLQKLAELNHQKLESNKIPELDMTEVAVPFMANIPILLYPDGPR 180
DB 132 -PKTICWNLGLCQPRQ-----PD----- 148
QY 181 SKPQPKNGDVQCQDCIQMTVDIQTAVRTNSTFVALVEHVKECDRLPGMGADICKNYIS 240
DB 149 --PQPE-----PGMPSPLPNSLS 164
QY 241 QYSEIATQMNMHQPKEICALVGFCDVEKMPQTLVPKAVY-ASKNVITPALVELVPIKKA 299
DB 165 ETSP-----DLVPPRLPRALSVWPGPH--TQDLSEQ 194
QY 300 EVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKELD-AFDKMSCKLPKSLSECEVVDVY 358
DB 195 QFPIPLP-YCRLCKTLKRVQAMT-----PKGYLAAMAVQCHVPLVAGTICQCLAERY 248
QY 359 GSSILSLILEVSPELVCSMLHLS-----GTRLPALTVAHTQPKDGFCEVCKLGVYL 413
DB 249 TVLLDLALSHLRLPOLVCGVLVLRCSMDNSAGLVLPALSLSEGWLPODSECDLCLSVT--- 305
QY 414 DRNLKSTKQETLALEKSGCSFLPDYQKO-CDQFAVEYEPVLIIELVEMDPSFVCLK 472
DB 306 --TRAWNSSELARQAMHQAAC--LSSPLDRQCKQVEQYTTPOLALDEPRGDPRTTQCA 361
QY 473 IGACPSAHKPL 483

```

DB 362 LGVCRGANPL 372

Search completed: May 5, 2004, 13:33:41
Job time : 52.6524 secs